IN THE UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

RELX INC.,

Plaintiff,

vs CASE NO: 16-CV-9718 (AKH)

INFORMATICA LLC,

Defendant.

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

INFORMATICA LLC,

Counter-Plaintiff,

vs

RELX INC., RELX GROUP PLC,

Counter-Defendants.

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

VIDEOTAPED DEPOSITION OF EXPERT

CHRISTOPHER RUCINSKI

June 20, 2018

9:52 a.m.

One International Place

Boston, Massachusetts

Deborah J. Bateman, Court Reporter

APPEARANCES OF COUNSEL

On Behalf of the Plaintiff RELX Inc.:

ROBERT J. SCOTT, ESQ.
BARBARA FREDERIKSEN-CROSS, ESQ.
SCOTT & SCOTT, LLP
550 Reserve Street, Suite 200
Southlake, Texas 76092
214.999.2905
Rjscott@scottandscottllp.com
Julie@scottandscottllp.com

On Behalf of the Defendant Informatica LLC:

SCOTT W. DOYLE, ESQ.
Greenberg Traurig
2101 L Street, N.W., Suite 1000
Washington, DC 20037
202.331.3143
Doyles@gtlaw.com

Also Present:

Kayli Smendec, Summer Associate
Couirey Eckmayer, Videographer
Paul Levy, Vice President, Informatica
Barbara Frederiksen-Cross

INDEX OF EXAMINATION

INDEX TO EXHIBITS

DEPOSITION OF CHRISTOPHER RUCINSKI

June 20, 2018

THE VIDEOGRAPHER: This is Tape No. 1 to the videotaped deposition of Christopher Rucinski in the matter of RELX Inc., plaintiff versus Informatica LLC, defendant, being heard before the U.S. District Court, Southern District of New York, Case No. 16-CV-9718 (AKH).

This deposition is being held at Greenberg Traurig, LLP, in Boston, Massachusetts. Today's date is 2018, June 20, and the time on the record is 9:51 a.m.

My name is Couirey Eckmayer, and I am the videographer. The court reporter is Deborah Bateman.

Counsel, will you please introduce yourselves and affiliations, and the witness will be sworn in.

MR. DOYLE: Scott Doyle, counsel to Informatica Corporation.

MR. LEVY: Paul Levy, Vice President for Informatica.

MS. SMENDEC: Kayli Smendec, summer associate for Greenberg Traurig with Informatica.

MR. SCOTT: Rob Scott with Scott & Scott representing the RELX parties.

MS. MACHAL-FULKS: Julie Machal-Fulks also

representing the RELX parties with Scott & Scott.

MS. FREDERIKSEN-CROSS: Barbara Frederiksen-Cross. I'm here as an assistant to counsel for Informatica.

CHRISTOPHER RUCINSKI, having been first satisfactorily identified and duly sworn, testified as follows:

EXAMINATION

BY MR. DOYLE:

Q. Good morning, Mr. Rucinski.

A. Good morning.

Q. Is that how to say your name properly, "Rucinski"?

A. Yes, that's correct.

Q. Okay. You understand today you're under oath?

A. I do.

Q. And do you understood that verbal answers are necessary?

A. Yes.

Q. Okay. One thing that I ask is I'll be asking questions, and I would appreciate if you allowed me to get my question out onto the record and then you can

provide your answer, and I'll try not to interrupt your answer with a new question. Do you understand that?

A. That makes sense to me.

Q. And do you understand that you need to answer despite objections unless instructed otherwise by your counsel?

A. Yes.

Q. And today I'm going to assume you understand any question you answer. Is that fair?

A. Sure. And if I don't understand it, I'll just ask you to --

Q. Okay.

A. -- clarify it somehow.

Q. Great. And if you need a break today, just let us know. All I ask is that if we have a question on the record, that you answer that first. Is that all right?

A. Okay.

Q. As you sit here today, any reason why you cannot provide full and truthful testimony?

A. As I sit here right now, I can't think of any other reason.

Q. Okay. Mr. Rucinski, what did you do today to prepare for your deposition?

A. Today I woke up, I reviewed my two reports briefly, and then met with counsel briefly for about ten minutes before walking up here. And that's all I did this morning.

Q. Well, other than today, what did you do to prepare for your deposition?

A. Yesterday I reviewed my reports in more detail. I also met with counsel for about two hours. I reviewed the deposition transcript of Ms. Frederiksen-Cross. And in total, yesterday, apart from meeting with counsel, I probably spent three to four hours preparing.

Q. What about before yesterday?

A. Well, I reviewed the deposition transcript of Ms. Frederiksen-Cross before yesterday, but that's all I can think of right now. I only prepared in the last few days.

Q. Did you have any other meetings with counsel in preparation for your deposition other than the one you had today and yesterday?

A. No.

Q. Did you have conversations with anyone else in preparation for your deposition other than counsel?

A. So let me amend. So yesterday I spoke on the

phone with Dwight Groff to get some clarity on an issue that was raised during Ms. Frederiksen-Cross's deposition, and that call was for about ten minutes or so.

Q. And what was that issue?

A. The issue was the -- in one of the tabs on a spreadsheet relating to the utilization of the RELX servers, there was reference to a user with a name of, I think it was, 1UInform, and that was raised during the deposition of Ms. Frederiksen-Cross. And so I wanted to talk to Dwight to get some better clarity on -- on why that name had showed up in that spreadsheet.

Q. And what does that name indicate?

A. Understanding from talking with Mr. Groff is that that is the -- the name of the user process that was set up when the servers and the ICCE grid -- and ICCE, that's spelled I-C-C-E -- when those servers were set up with the Informatica software, that was the user that was set up to run parts of the Informatica software, and so that's why it bore a name resembling Informatica.

Q. So does that mean that content data was still going across the Informatica platform at that time and date?

MR. SCOTT: Objection.

A. Would you clarify the time and day that you mean in your question?

Q. Yeah. Associated with the spreadsheet you were looking at.

A. It's my understanding from talking with Mr. Groff that that user name was associated with -- or was the user name associated with all of the processing that appeared on the spreadsheet; however, after around November 7, that name -- or that user name was still present in the system although it was executing software that was not Informatica software.

Q. And what software was it executing?

A. So from talking with Mr. Groff, my understanding is that it was executing other parts of the ICCE platform which has since, and around that time period, changed to not incorporate the Informatica software.

Q. Do you know if any of the Informatica software was still being used on November 17, 2017?

A. Specifically on that date, my understanding from Mr. Groff's sworn testimony is that it was not running at that time.

Q. Did he provide you any more reports to look at after that date and time?

MR. SCOTT: Objection.

A. After the date and time you mentioned, I don't recall seeing any other reports apart from the ones I mentioned in my report.

Q. Did you ask him to see more reports after November 17?

A. I didn't. Because, as far as I understood, the Informatica software wasn't running when that date came to pass, so it didn't seem relevant to the items I was looking at for my reports.

Q. So is it your -- it's -- your understanding is that after that date, no more data was running through the Informatica platform?

A. The way I would state it is that the Informatica software -- well, let me restate that.

On that -- as of that date, the Informatica software wasn't being used in the ICCE platform to process files; so while the ICCE platform was processing files, it was no longer using Informatica software at that time.

Q. Was Informatica software being used for any other -- as part of any other platform other than ICCE after that time?

A. As far as I know, it was not.

Q. Who told you that?

A. I believe Mr. Groff mentioned that to me. I think it may also be in the e-mail that he sent, but I'd have to take a second look at that to be sure.

Q. And did you independently verify that?

A. Well, having spoken with Mr. Groff and having read his deposition testimony, I didn't -- because there was nothing to suggest that it wasn't running at that time, I didn't feel it was necessary to look into the issue any further.

Q. What was the activity that, I think you called it -- was it U1 -- let me get this -- U1Inform, is that the name?

A. That's my recollection.

Q. Okay. What was the activity that U1Inform was processing prior to the date November 17 in 2017?

MR. SCOTT: Objection.

A. So in general, the -- that user process was running processes associated with the ICCE platform. Some of those processes were related to the Informatica software, but others were related to other parts of the ICCE platform. And, then, in general, after the ICCE platform ceased using Informatica software, it was only processing other programs that were -- or processes that

were part of the ICCE platform.

Q. Which processes were associated with the Informatica software?

A. So there were a number of workflows that were, we'll say, kind of defined by the Informatica software in terms of their overall structure, and then there were specific data transformation processes that were executed as part of the Informatica software.

Q. Do you know if that term U1 -- I'm sorry -- 1UInform is still used as part of the ICCE platform?

MR. SCOTT: Objection.

A. I think it was U1Inform, but I'd have to check the document --

Q. Okay.

A. -- to make sure. And sorry, would you repeat that question?

Q. Yeah. The question is do you know if that indicator is still used as part of the Informatica -- I'm sorry. Strike that.

Do you know if the indicator U1Inform is still used in the ICCE platform today?

MR. SCOTT: Objection.

A. I'm not sure if it is or not.

Q. Does Inform -- the -- so the name U1Inform, is

the "inform" portion of that name, does that mean Informatica?

MR. SCOTT: Objection.

A. I'm not sure what it means per se, but my understanding from talking to Mr. Groff is that it was related to the fact that when the ICCE platform was first developed, it was designed to use Informatica software.

Q. Which parts of the U1Inform activity were Informatica before November 2017?

MR. SCOTT: Objection.

A. Would you clarify which dates you're referring to?

Q. Yeah. I'm sorry about that. Which parts of the U1Inform activity were Informatica before November 17, 2017?

MR. SCOTT: Objection.

A. My understanding is that -- first of all, with respect to the dates, my understanding is that the Informatica portion of the ICCE platform was -- no longer took in any documents after November 7, 2017.

Q. Okay.

A. Not the 17th.

Q. Thank you.

A. But on that date and prior, I think as I

already discussed a little bit, the portions of the ICCE platform that were Informatica were related to, for instance, the workflows that Informatica defined, and specifically, the data transformation processes that were executed.

Q. And what workflows were those?

A. There were a number of workflows related to processing documents related to, for instance, first retrieving the document from the provider where it originated, and then converting it into a text form, and then doing some other transformations to get it into a master form.

And to be clear: There were portions of those workflows that were implemented by RELX, and so those -- those -- the execution of those workflows were more in line with RELX customized software. There were specific workflows that were implemented using Informatica software. And then the workflows themselves were kind of part of the framework that Informatica provided.

Q. What percentage of the workflows were associated with Informatica?

A. Would you clarify the time period that you're talking about?

Q. Prior to November 7, 1918 [sic].

A. My understanding is that that changed over time. The ICCE platform went through at least a couple major versions, and I think there was at least one minor version in between. So in general, my understanding is that the ICCE platform originally started out using more Informatica workflows that were implemented by Informatica, and then over time, transitioned to using fewer of those workflows.

Q. And when you say "started out," what -- approximately what year was that?

A. My understanding is that the ICCE platform began processing documents using the Informatica software for part of that in 2012.

Q. And so in 2012, approximately, what percentage was Informatica of the workflows?

A. Of the workflows that were implemented by Informatica, I'm not certain, as I sit here. There was a diagram in my report where I labeled a number of -- a number of workflows that were implemented, for instance, in Java that RELX created themselves, and in Perl that RELX created themselves, and then there were others that were labeled as Informatica specific.

Q. Okay. What about with respect to the data transformation processes? Were those all Informatica?

A. So when I say data transformation, I'm specifically referring to the Informatica data transformation software. So yes, by data transformation, you know, upper case, shall we say, I'm talking about Informatica software in that instance.

Q. So before November 7, 2017, is all the CPU data reported in the charts ICCE related?

MR. SCOTT: Objection.

A. Would -- would you state that one more time?

Q. Yeah. Prior to November 7, 2017, is all the CPU data and CPUs reported in the charts ICCE related?

MR. SCOTT: Objection.

A. My understanding is that because the -- well, let me state it this way: The CPU data was reported from the operating systems that were executing on the servers that were part of the ICCE platform. So it's my understanding that they were -- those CPUs were being used to run the various components of the ICCE platform. I'm not aware, as I sit here right now, of other things they might be doing that were apart from the ICCE platform, and, again, noting that the ICCE platform included parts beyond the Informatica software.

Q. Are you aware that some of the U1Inform numbers started on May 15, 2017?

A. That would be consistent with my understanding because the U1Inform user, as far as I know, was the user that was executing processes on the servers in general.

Q. And what did those numbers relate to --

MR. SCOTT: Objection.

Q. -- the new numbers started on May 15?

MR. SCOTT: Objection.

A. Would you clarify what you mean by the "numbers"?

Q. What did -- what did they identify?

MR. SCOTT: Objection.

A. Are we talking about the utilization numbers for the servers or something different?

Q. Certain U1Inform numbers started on May 15, right?

MR. SCOTT: Objection.

A. I'm still not sure what you mean by "numbers" in this instance.

Q. Okay. Some of the U1Inform identifiers were started on May 15 --

MR. SCOTT: Objection.

Q. -- 2017, are you aware of that?

A. I'm sorry. I still don't understand what you mean by -- which numbers are we talking about?

Q. Is it fair to say that prior to November 7, 2017, that some of the U1Inform were ICCE?

MR. SCOTT: Objection.

A. When you "some of the U1Inform," I'm not sure what you mean by "some of the U1Inform."

Q. Well, I mean -- yeah, let me specify. U1Inform indicated -- you testified earlier, right, that U1Inform referred to Informatica processing; is that right?

MR. SCOTT: Objection.

A. U1Inform is the name of a user on a system. My understanding is that it was named initially because when the ICCE platform was being developed, it was being developed with the incorporation of the Informatica software in mind.

Q. Okay. And, then, are you -- before November 17, 2017 -- strike that.

What documents did you review for this deposition?

A. To prepare for this deposition, I reviewed my own reports, two of them; I also reviewed parts of the deposition of Nalin Mishra; and I reviewed the transcript of Ms. Frederiksen-Cross's deposition.

Q. Which portions of the Nalin Mishra deposition

transcript did you review?

A. As I sit here right now, the portion I remember is the one in which he was testifying about an e-mail that he received related to the installation of Informatica software onto servers comprising 104 cores and how he -- he was aware that it was installed on servers with 104 cores.

Q. And who was that e-mail from?

A. As I sit here right now, I don't remember who it was from.

Q. Was it from Charles Sedlacko?

A. I think I'd have to see the deposition transcript again to refresh my memory.

Q. Do you recall approximately the date of that e-mail?

A. I think it -- as I sit here right now, I think it was around 2013.

Q. And what was in the e-mail?

A. I think it was related to what -- what servers had the Informatica software installed and how many cores those servers had.

Q. So was it your understanding in 2013 that the software was deployed on 104 CPUs?

A. The way I would state it is that at some point

in 2013, Nalin Mishra installed the software on servers, and those servers, in total, had 104 cores at their disposal.

Q. And do you know if Informatica software was installed on those servers?

A. I believe Informatica software was installed on the servers that were running the ICCE platform because my understanding is that the ICCE platform took advantage of the Informatica software for portions of time in order to perform its function.

Q. All right. And those servers that Informatica software was installed in had 104 total CPU cores, correct?

A. Those servers that Informatica software was installed on had 104 cores to execute the ICCE platform, including the Informatica software, for certain periods of time.

Q. And at that time, how many licenses did RELX have in terms of CPU cores?

A. Would you clarify which time in particular you're talking about?

Q. Same time we're talking about here. You said 2013.

A. Right. I'm asking for your question,

specifically at what point in 2013 are we talking about.

Q. The point of this e-mail that you reviewed.

A. I don't recall the exact date of the e-mail. Would you help me to understand which date you're talking about in particular?

Q. That's the date, the date of the e-mail.

A. I don't --

Q. We'll pull it out later if you don't -- if you don't recall. Do you not recall?

A. I don't recall the date of the e-mail.

Q. Okay. Any other documents that you used to prepare for this deposition?

A. Would you state that one more time?

Q. Any other documents that you looked at to prepare for this deposition?

A. There was one other document that I received from counsel. It was sort of a generic sort of advice for how to prepare for a deposition. It was two pages long.

Q. Going back to what you reviewed in the Nalin Mishra transcript, was it your understanding in 2013 that when the Informatica software was deployed on 104 cores, that RELX was over deployed in terms of its license agreement?

A. The way I would state it is at some point in 2000, the Informatica software was installed on servers that had 104 cores at their disposal to execute the ICCE platform including the Informatica software, and that I believe at some point in 2013, the agreement between RELX and Informatica provided RELX with licenses for 72 cores.

Q. And so would that be an overdeployment?

MR. SCOTT: Objection.

A. My understanding is that the -- to the extent there's a legal issue here, that's not what I'm here to testify about, but it is true that 104 cores is more than 72 cores, if that's your question.

Q. Anything else that you did to review for this deposition?

A. As I sit here right now, that's all I can think of.

Q. Okay. How long have you been employed at Stroz?

A. So that's a bit complicated because Stroz Friedberg acquired a company called Elysium Digital in 2015, so that acquisition happened around July of 2015. So at Stroz, I've been employed for about three years. Prior to that at Elysium, where I was doing the same

sort of work, I was there for about five years before that.

Q. Okay. What was your first job after graduating college?

A. Working at Elysium was my first job after graduating college.

Q. Okay. So you've been employed for eight years since college?

A. Yes.

Q. How would you characterize your expertise in computer science?

A. Well, in general -- so I have a degree from Princeton University in computer science. I also received a certification, the title is the GIAC Certified Forensic Examiner Certification, that was in 2015. Over my employment at Elysium and Stroz, I've looked at software and technical documents for around 80 different matters including some copyright cases.

Q. What's the GE Forensic -- did I say that correctly?

A. It's G -- GCFE.

Q. GCFE?

A. Yep.

Q. What is that?

A. So it stands for GIAC Certified Forensic Examiner. GIAC, I believe, stands for Global Information Assurance Certificate. They're a sort of industry body that provides training for different certifications related to the forensic examination of computer artifacts, and they have other certifications for security and items of that nature.

Q. And do you still do that work today?

A. Would you define "that work"?

Q. Anything covered by your description of forensic examination.

A. Well, I think "forensic examination" is a fairly broad term. I would in -- in my mind, it means anything related to artifacts that come from computers, so that's most of the work that I do.

Q. Okay.

A. Maybe all of it.

Q. When was the first case you were retained as an expert in?

A. That would be Shurtape v 3M, and that was in 2013.

Q. And when was the last time you were deposed?

A. That was sometime in 2015 for the BMG v Cox matter.

Q. Do you know if the list that you provided as part of your expert report is exhaustive with all of your engagements?

A. It includes all of the engagements that I either submitted an expert report for, that I was deposed for, that I testified at trial for, that I submitted a declaration for. And with those qualifiers, yes, it's exhaustive.

Q. Okay. And how many copyright infringement cases have you been retained as an expert?

A. So including this case, there are two cases, the other case being BMG v Cox.

Q. I'm sorry. BMG what?

A. BMG v Cox. That's Cox Communications, the Internet service provider.

Q. Okay. Who was plaintiff in that matter?

A. BMG was the plaintiff.

Q. And who were you retained by?

A. I was retained by the law firm Fenwick & West, and they were representing Cox Communications.

Q. Did you testify at trial?

A. I did.

Q. And did it involve copyright infringement software?

A. That case was about copyright infringe- -- excuse me -- copyright infringement related to sound recordings, and there was a software issue related to that that I was retained for.

Q. And what was the software issue?

A. So the -- in that case, the software issue was that there was a third party to the case called Rightscorp, and Rightscorp had developed software that they claimed was able to monitor computers who were using the -- or computers that were using the BitTorrent protocol to exchange data related to files. And so I was retained to examine that software.

Q. And what was your conclusion with respect to your examination?

A. The case was a long time ago. I think I would want to refer back to my expert reports for that matter before I stated, or summarized in any way, my conclusions.

Q. What year was it?

A. It was 2015.

Q. Do you have any recollection of what you actually did?

A. Sure. So the Rightscorp software spanned a number of Java files, and so I reviewed those Java files

to get a -- an understanding of how they executed. I also looked at -- there were a number of database records that were also produced in that matter that were related to recordkeeping for how the Rightscorp software kept track of, for instance, which IP addresses it was monitoring, which sound recordings it was trying to find. And it also kept track of, I think, just other kind of recordkeeping that was related to the operation of the software.

I also reviewed deposition transcripts of employees from Rightscorp who testified about how the software operated. I also attended a couple of those depositions to better understand how the software operated.

Q. Did you testify that the software did its job in terms of monitoring?

A. Well, my recollection, as I sit here right now, is that there were certain things that the software didn't do and that there were issues related to recordkeeping of the software itself, which is to say that rather than using a version control system to keep track of multiple versions over the time period that the software was executing, there was no such system in place, and so there was no dispositive record of what

software was running at certain times in the past when the software was alleged to have correctly created records related to how data files were alleged to be shared with the BitTorrent protocol.

Q. Did Cox win the case?

A. My recollection is that BMG won the case initially, and then Cox appealed the case, and I believe it's going to retrial sometime soon.

Q. Do you know what issues are going to retrial?

A. As I sit here right now, I'm not sure exactly what issues are going to retrial.

Q. Do you know if any of your opinion testimony was excluded by the Court?

A. My understanding is that none of it was excluded.

Q. Have you ever been involved in copyright cases where the allegations were directed to executable code?

MR. SCOTT: Objection.

A. Yes, several.

Q. Did you testify in any of those cases?

A. I don't think I did for any of those cases.

Q. Were you deposed in any of those cases?

A. No, I was not for any of those.

Q. Did you draft expert reports for any of those cases?

A. For those cases related to?

Q. Executable code.

A. Well, related to copyright for executable code --

Q. Yeah.

A. -- for those cases. Other than the ones -- well, for the cases we're talking about where there was issues of copyright related to executable software, for none of those cases did I write an expert report.

Q. And how many cases were there that you were retained on?

A. I'm not sure, as I sit here right now, but I would guess maybe ten or so.

Q. Can you recall what they were?

A. Well, let me see. So there was a case recently that was related to -- it was software for keeping track of wedding venues in the Las Vegas area and in general, and so in that case, I looked at two different implementations of the -- of that software, and there were copyright claims with respect to whether certain code was copied by the defendants.

There was another case that I was recently

retained as an expert related to a dispute between a software vendor and a cosmetics company, and the software vendor had been retained to help the company to implement part of their online kind of magazine spread, and then the company had retained another vendor afterwards. And the allegation was that there -- there was still copyrighted code in the system from the previous vendor.

Those are the ones that come to mind right now. I think there were probably others because those were both in the last few months, and we do frequently get retained for copyright -- for cases where there's a copyright issue related to software, usually a company's trade secret claims for software.

Q. With respect to those two cases, who were you representing?

MR. SCOTT: Objection.

A. Do you mean in terms of plaintiff or defendant or the specific parties that retained us?

Q. Those specific parties that retained you.

A. So generally, we try to not disclose who we are retained by unless it is somehow released to the other side, so let me just make sure that's the case. So for the cosmetics company, the other side does know about that, so we were retained by Mary Kay. They -- and they

are the defendants in that matter. That's in state court in Texas, I think.

And then, the other case, we were -- or I was retained by the plaintiffs. And there were a number of companies that were named as parties, one of them was Chapel of the Flowers. That was a company in Las Vegas. I think there was another corporate entity in there somewhere, but there was a merger and so I forget who the other company was.

Q. Who was the plaintiff that retained you in that case?

A. Chapel of the Flowers was the company.

Q. In those cases, or any others, did you actually examine the executable code?

A. Well, in those cases, I was looking at the source code --

Q. Okay.

A. -- as distinguished from the executable code. The source code is the code that's human readable and is usually what -- it's my understanding, at least, that tends to be the more important thing to look at with respect to copyright cases of software, if you want to compare, you know, someone's work, a party's expression of -- of something.

Q. And what analysis did you perform in those cases?

A. Well, I reviewed the -- the source code for both of the parties and looked for -- well, I need to distinguish between the two of them.

The case for Mary Kay is quite recent, so for that case, I looked at a limited amount of code that is available -- and by "code," I mean source code. So there's some source code that's available on the Mary Kay website. There's a small amount that has been produced in the litigation. I don't think any of it I've seen so far has been confidential to the plaintiffs in that matter.

For the case involving Chapel of the Flowers, there was production from -- from both sides, and so I reviewed the source code implementations for both the plaintiffs and defendants and looked at previous versions of the source code for plaintiffs and defendants, and tried to see to what extent there was sort of exact matches in the literal expression in the code, and also whether there was structural similarities between the two implementations.

Q. And is that your understanding of what you need to do as part of copyright infringement: compare the

plaintiffs code to the defendant's code?

MR. SCOTT: Objection.

A. Well, I can't speak to it in general because I'm not a lawyer, but for those specific cases, those seemed to be the most relevant questions and the ones that I was asked to -- to examine.

Q. So for those two cases, you examined the source code of your client and compared it to the source code of the -- of the other party?

A. Well, it's not true for the Mary Kay matter because I didn't have access to the other party's code.

Q. Okay.

A. That's just from public -- what's available on their website. And then in the matter for Chapel of the Flowers, yes, I did look at the source code for both the plaintiffs and defendants and tried to see to what degree they were similar.

Q. And was there any determination of copyright infringement in that case?

A. Neither of those cases have even got through expert reports at this point.

Q. In either of those cases -- well, not Mary Kay, but the other one, Chapel of the Flowers, do you believe there to have been copyright infringement?

MR. SCOTT: Objection.

A. Well, I haven't come to an opinion yet in either of those cases.

Q. If it turns out that there's a substantial similarity between the code that you're examining and comparing, would you come to the conclusion that copyright infringement occurred?

MR. SCOTT: Objection.

A. It's hard to say because there's a lot of factors. So, for instance, when you're comparing source code implementations, just because there is some similarity, doesn't necessarily convince me that there is a copyright violation. Because, for instance, both implementations of the source code could have drawn from a third source that could perhaps be open source. There could be also other reasons in the case why it was, for some reason, okay for a certain code to be used, if there was some other agreement, for instance.

And in general, I don't -- I wouldn't want to go on the record as saying in a hypothetic universe, if there were similarities, that I would always come to a certain conclusion. I just think there are other factors that might influence my opinion.

Q. Okay. Did you do that in this case?

MR. SCOTT: Objection.

A. Would you define what you mean by "do that"?

Q. Compare source code to source code.

A. So in this case, my understanding is that RELX never had access to Informatica source code, and I didn't review it because there wasn't access. And I understand that Ms. Frederiksen-Cross also has not reviewed the source code.

Q. Did you ask for source code in this case?

A. I don't think I ever explicitly asked for it in this matter.

Q. Do you believe there's a question as to whether or not code was copied in this case?

MR. SCOTT: Objection.

A. When you say "code," do you mean source code or do you mean something different?

Q. Source or executable code.

A. So in the previous cases, just to clarify, we've been talking about with Mary Kay and Chapel of the Flowers, that's -- that's all been source code. In this matter, to the extent there was any copying, I believe it's only executable code. And I think there are -- there are questions that are in dispute about, for instance, who -- or what party was responsible for

installing the Informatica software, where installing would include some copying onto the machines in the ICCE platform, for instance.

Q. But what did you do to satisfy yourself that copying either occurred here or did not occur here as it related to the copyrighted software?

MR. SCOTT: Objection.

A. Well, I understand that Informatica software was copied onto the machines, and I don't think it was -- well, there wasn't much of a question there from a computer science perspective in terms of who was actually responsible. That seems more like a question with respect to contractual obligations or who actually performed certain actions.

I did, however, review some of the installation scripts that are mentioned in my report that relate to the installation of the software.

Q. So what's your understanding of the copying in this case?

MR. SCOTT: Objection.

A. It's a bit of a broad question. Could you maybe narrow it a little bit so I can give a direct answer?

Q. Well, does the -- when you load or install the

software, is it copied on the servers?

A. Well, maybe we can do it one at a time. You said load and install. Which --

Q. Go ahead and answer each. Start with load.

A. Would you define what you mean by "load" in that case.

Q. Do you not know what "load" means?

A. I'd like to know what you mean in terms of the question because I don't understand --

Q. Do you have an understanding of what the term "load" means as it relates to software?

A. Well, the way I would phrase it is like this: When software executes on a computer, generally the software is resident on persistent storage on the computer, so that could be a hard drive, for instance. And then when the program is executed by the computer, portions of that program are transferred to memory as necessary so that the processor can access the instructions it needs to access and relevant data perhaps. Excuse me. And so in that sense of loading where it's transferring from a hard disc to memory, you have portions of the program being transferred in the normal execution of software on the computer.

Q. So are they copied in that case?

MR. SCOTT: Objection.

A. Again, the way I would state it is there are portions of the program that exist at the same time on the hard disc as well as memory -- RAM, in this case -- while the processor is executing instructions, and then different portions may be swapped in or out of the memory as the execution of the program continues.

Q. And is it your understanding under copyright that that is a copy --

MR. SCOTT: Objection.

Q. -- that function?

A. That sounds like a legal question to me, and I'm not a lawyer, so I don't know if I can offer an opinion directly related to copyright.

Q. Well, I'm just asking -- you know, you've done this in other cases purportedly -- aren't you looking to see whether software gets copied --

MR. SCOTT: Objection.

Q. -- onto the computers? Isn't that what you're doing in these other cases?

MR. SCOTT: Objection.

A. In the other cases, I'm -- I'm more looking at -- for a record of the source code that both of the parties possess, is there an expression that is similar

or the same between them. It's -- the questions in those cases haven't turned on whether or not certain software was actually on a computer. It's just about the software that's -- that's produced.

Q. Do you believe in this case that installing software, the Informatica software makes a copy of the software on the computer?

A. I would clarify that if you're -- if you're installing the software on a computer, you are placing another copy of the executable form of that software onto the computer on which you're installing it if you're installing to the hard drive.

Q. Okay. And each time you install, the software makes a copy?

MR. SCOTT: Objection.

A. I wouldn't say that's always the case.

Q. In the case here with Informatica, do you have any instance that you can recall where the Informatica software was installed and a copy wasn't made?

MR. SCOTT: Objection.

A. I'm sorry. Would you restate that one more time?

MR. DOYLE: Can you read that back, please.

(Question read)

A. So the case I examined most directly was the upgrade of Informatica software to Version 9.6.1. And while there was -- while there were certain, at least, portions of the software that was rsynced -- and I'll spell that. Rsync is r-s-y-n-k -- or sorry -- y-n-c -- in that case, there were at least portions of software that were transferred onto the servers. But having not looked in detail in terms of whether a previous version was still resident, I'm not -- basically, I'm not sure whether there was a complete copy of the new version that was transferred on or if it was just an incremental improvement to the existing software on those computers.

Q. In general, when you execute software, does that also make a copy of the software?

MR. SCOTT: Objection.

A. So I'll have to explain like I did earlier. When you execute a program on a computer, what generally happens is -- assuming that the executable software is distinct from the source code, is resident on the hard disc. Portions of that executable software are copied into memory so that the processor can execute it more easily.

Over time, other portions of the software may

be copied into memory and certain portions may be removed from memory to make space because it's sometimes the case that there isn't enough memory available on the computer to hold the entire program, and in general, a computer won't copy an entire software program into memory just to execute a small part of it.

Q. So here, executing the Informatica software makes a copy of the software?

MR. SCOTT: Objection.

A. Well, I'm trying to explain the -- the complexity of it. I certainly wouldn't say that the entirety of the Informatica software was copied into memory without being provided certain evidence that would suggest that. Because, in general, when you have a program that's resident on hard disc, only a portion is copied into memory in order to execute that specific portion of it.

Q. But would you expect at least portions of the Informatica software to be copied when you execute the software?

A. Well, it's my understanding that a computer processor can't execute, or in general, doesn't execute programs directly from the hard disc. So in order to execute portions of the program, yeah, I think portions

would have to get moved into memory for the processor to execute it.

Q. Do you know what portions of the Informatica software were copied here?

MR. SCOTT: Objection.

A. Well, I don't have a record of exactly which portions were copied. In general, if -- if there is an instruction for a program that is going to be executed by the processor, it needs to get moved into memory. So if there are portions of a program that either -- that are part of the executable but in the course of the program wouldn't actually execute at any point in time, either because that -- that executable code just logically with the program wouldn't ever execute, or if it was configured to never execute, or if the circumstances for that execution never arose, then those certain portions wouldn't get copied into memory because the processor wouldn't have to execute them.

Q. But some portions are copied into memory, correct?

A. During the execution of a software program, certain portions of the software program do need to be moved into memory if they are to be executed by the processor.

Q. Are there separate copies made of the software when the software is run across multiple cores in a server?

MR. SCOTT: Objection.

A. I think it would depend on the software. I can give you an example. So if you have a program that is -- those running on multiple threads -- so a thread is kind of like a -- it's like an individual task that might get executed as part of a program, and that individual task can run on multiple CPUs, for instance -- but all of those tasks, if they're run as part of the same program, might share the same it's called "address space" for memory. Computers have a limited amount of memory, and programs have their own address space to kind of make sure that they don't overwrite other portions of the memory that other programs are using.

So if the program has multiple threads, and they're each running on different processors, they may all access the same address space, which means that they don't actually need to have multiple copies of the software in memory. They can just all work off of the same one because they're all accessing the same address space as the memory.

Q. And was that -- what is your understanding

of the situation with respect to the Informatica software?

MR. SCOTT: Objection.

A. I'm not exactly sure how the Informatica software works in terms of whether it has multiple threads or whether it's implemented in some different way.

Q. Did you ask for that information?

A. I don't recall that I did at any point.

Q. Why not?

A. Generally, the reason I was retained here was to look at primarily the question of the degree to which, if any, RELX benefited from having more cores available for the Informatica software that was part of the ICCE platform -- at certain times, more cores than the number of licenses that they had at the same time.

On the question of copyright, I was -- the primary thing I was looking at was the upgrade scripts and, in general, how portions of the Informatica software were -- were installed on the disc and then, you know, its kind of normal operation.

Q. So if you're called to testify at trial and you're -- you're asked did copyright infringement occur with respect to the Informatica software, what

would be your testimony?

MR. SCOTT: Objection.

A. Well, I don't think I can testify directly on the issue of copyright infringement as that strikes me as a legal conclusion.

Q. So you won't provide any testimony to that question?

MR. SCOTT: Objection.

A. I may provide testimony that counsel might rely upon to come to legal conclusions, but I myself am not here to offer legal opinions.

Q. Okay. What testimony would you provide that counsel may rely on?

A. Well, my opinions are laid out in my expert reports, and I took a fair bit of time to -- to write them in a careful way and one that was correct, so I hesitate to restate those opinions as I sit here right now lest I forget something that I wrote down.

Q. So are you saying you need your expert reports at trial to testify if you're asked that question?

A. That's not what I'm saying. I'm saying, as I sit here right now, if you're asking me what opinions am I going to give, then my opinions are in my expert reports. I just don't want to misstate them.

Q. So what are your opinions with respect to whether or not copies were made and whether or not there's copyright infringement?

MR. SCOTT: Objection.

A. So I'm not going to speak to -- I'm not going to speak to copyright infringement because that's a legal question, but with respect to opinions related to question of copyright, as I said, my opinions are stated very carefully in my expert report, and I wouldn't want to misstate them here as that would be a disservice to everyone here.

Q. So you're not -- are you willing to summarize your opinions as it relates to copyright infringement?

A. I can sit here and, from memory, try and recall at a high level what those opinions are, if you'd like.

Q. Please do.

A. So this will be a non-exhaustive list as I would prefer to have my expert report in front of me, but, in general, the opinions that I offer in my report relate to whether or not copying that occurred included source code or whether instead it was related to executable code. And my opinion is that the copying that had occurred at certain points in time and certain

portions was related to the executable code and not to source code.

Q. So there was copying of executable code in this case; is that correct?

A. There was at least portions of executable code that were copied at specific points in time, as far as I understand.

Q. And what are those specific points in time?

A. So, as I sit here right now, my recollection is that executable -- portions of executable code were installed onto the servers of the ICCE platform as part of, for instance, the upgrade to Informatica Version 9.6.1. It's my understanding that for previous versions, executable code would have had to be on those computers as well and transferred from some other location.

There was also the issue of whether executable code was copied into memory during normal execution. And as we've already discussed, there are portions of executable code that will necessarily be copied into memory so that the processor can execute it.

Q. Do you know what portions of the executable code were actually copied into memory --

MR. SCOTT: Objection.

Q. -- in this case?

MR. SCOTT: Asked and answered.

A. So I think we already talked about that, but I don't have any records of precisely which portions of the executable code were copied into memory, and I'm not sure which specific portions were or were not copied into memory.

Q. Earlier, did you testify that you didn't know about the Informatica use of the threads?

MR. SCOTT: Objection.

A. I think what I said is that I'm not certain how the Informatica software is implemented with respect to whether it creates threads that are -- that use the same address space or memory.

Q. Well, if you're not sure whether or not Informatica software creates threads in the same address space and memory, how can you evaluate the use of the cores?

MR. SCOTT: Objection.

A. Would you repeat that?

MR. DOYLE: Can you repeat that, please.

(Question read)

A. So I think you might be getting to one of the primary purposes of my report which was to examine the degree to which RELX benefited from having software

installed on machines that had access to cores that were a greater number than the number of cores granted them in the license agreement. The question of whether specific threads have access to the same address space doesn't seem related to that question to me.

Q. So whether or not threads have access to the same address space has nothing to do with the use of certain cores?

A. Well, it's a -- it's a different question. The use of -- the question of address space is related to whether or not there -- there is going to be more than one version of the software in the memory. The question of a CPU core execution, that's more related to how much processing they need to do in order to execute the program.

Q. So it doesn't matter whether -- whether or not that the software is creating threads to a same address space or to different address spaces in order for you to determine whether or not the CPUs are being used?

A. So, again, the question that I -- I was trying to focus on in my report was whether RELX benefited from having the software installed on servers that had access to more cores than the 72 that they had through license agreements. And in order to answer that question, what I

looked at was the degree to which the different CPUs were utilized -- or rather, the degree to which the servers that had the CPUs were utilized.

The question of -- I feel like you may have a different understanding of use as it relates to that question because my focus was on the utilization.

Q. Can you answer the question I asked?

A. Would you restate it?

Q. Sure.

MR. DOYLE: Can we restate that question.

(Question read)

A. Would you clarify what you mean by "the CPUs being used"? Over what time periods or to what degree?

Q. At any time.

A. Okay. Would you repeat that one more time?

Q. The entire question?

A. Yes.

MR. DOYLE: Can you repeat it again. I'm sorry.

(Question read)

A. All right. So I think those two questions are unrelated because one relates to whether or not there are -- there needs to be one copy of a program in memory versus more. And the question of CPU use, as you have

defined as the degree to which certain CPUs were used at any point in time, is more related to how much processing they needed to do. So I think those two questions are not related.

Q. Can you still answer my question, though?

MR. SCOTT: Objection.

Q. Whether the two questions are unrelated or not, can you answer the question that I asked you?

A. Well, I think those two issues are related so the --

Q. So is the answer "no"?

A. Well, let me finish.

I think if you answer one of those questions, you don't get any information related to the other one, so I don't think one depends on the other.

Q. Okay. I just want to get a clear answer on the record.

MR. DOYLE: Objection.

Q. Okay? You can say "yes," or you can say "no" to my -- my question.

MR. DOYLE: Would you repeat the question, sorry, one more time.

(Question read)

A. So with respect to those two questions, I

don't think it matters except maybe if you got some information related to how much processing was done or how much processing would need to be done based on the number of threads, but that doesn't seem to matter with respect to whether or not it was the same address space.

Q. Can you explain how the Informatica software uses CPU cores?

MR. SCOTT: Objection.

A. And by using CPU cores, do you mean, as you stated earlier, that any CPU cores are used at any point in time?

Q. Yeah.

A. Well, I don't know exactly how the Informatica software was executed on certain CPU cores over the entire time period. I just don't have records for exactly which cores were used to process specifically the Informatica software versus other components of the ICCE platform, for instance.

Q. I didn't ask about the entire time period. I'm just asking how the Informatica software accesses cores.

MR. SCOTT: Objection.

A. Okay. So help me to understand. If you're not asking about the entire time period, which time

period are you asking about?

Q. Any time period. Do you just have a general how it works?

A. I'm unclear in the distinction between any time period and the entire time period.

Q. Okay. Let's start with the entire time period.

A. Okay.

Q. Can you explain how Informatica software uses cores?

A. I don't have any records for -- for the time period about which cores were -- were used by the Informatica software as distinct from the ICCE platform in general.

Q. Okay. Do you have some information on some time periods?

A. Well, let me go over the information that I do have. That might help.

So as I mentioned in my reports, there is a record of over the time period of a few years for each hour of each day for each server in the ICCE platform, the average utilization across all of the CPU cores in that server for processes that were related to the ICCE platform.

So the things that are missing in that data are, for instance, which specific cores were used on the server as opposed to just average CPU utilization in general, as well as what -- what specifically was being executed on those CPU cores, because it could be processes that are Informatica processes in the ICCE platform, or it could be other processes that were executing as part of the ICCE platform. Those are the two main pieces of information that, as far as I know, there's no record of for the time period relevant to this case.

Q. So you looked at the average utilization, for example, over a bunch of cores in an hour. And what would that provide you in terms of CPU cores?

MR. SCOTT: Objection.

A. So specifically, it's not just a bunch. It's specifically the CPU cores that a specific server had access to. That's the level of granularity that I have.

And the way I approached the question of whether or not RELX benefited from having Informatica software installed on computers with 104 cores, for instance, versus the 72 that they were licensed for, was to look at what -- what do I think would have had

happened if they, instead of having 104 cores available, had 72 or 56 over that specific time period. And so I looked at a time period towards the end of the relevant time period of this case where RELX did, in fact, have only 56 cores available to the ICCE platform, including the Informatica software part of it, and then compared that to the times that didn't.

And even though the granularity in terms of time with respect to the data that was available was an hour, I understand from -- from an e-mail from Mr. Groff that -- that RELX didn't really care about processing anything faster than at least three hours -- and I imagine for other documents, it was -- it may have mattered less in terms of the actual time -- so that, you know, granularity, while -- while it could have been a narrower granularity than an hour, it seemed --

(Reporter clarification)

A. -- while -- while the granularity of the data could have been less than an hour, having it at an hour seemed to be sufficient to examine the question of -- of the benefit of having other cores available.

Q. Do you know if there's any information that could have told you which specific cores were used at any particular time?

A. I don't think there's any better information than the information in that spreadsheet that I was referring to in my previous answer.

Q. That's not what I asked. I said do you know whether or not there's any information that would identify which specific cores were being used.

MR. SCOTT: Objection. Asked and answered.

A. I think the question is do I know, and I'm unaware of other information that would provide clarity on that -- that question.

Q. Okay. Again -- and now backing away from that, I'm not asking which -- in this question, I'm not asking which cores were used, just how does it use the cores?

MR. SCOTT: Objection.

Q. I'm not asking for average utilization. I'm just asking how does the program use the cores --

MR. SCOTT: Objection.

Q. -- to do processing?

A. Well, in general, computer programs use CPU cores to execute instructions in order to execute the program. And there are some programs that can use more than one core at a time, maybe over different periods of time.

Q. Can Informatica use more than one core at a time?

A. Well, the best information I have about this suggests that under certain circumstances, Informatica software can use more than one core at a time.

Q. And what circumstances are those?

A. Well, I'm not sure about all the circumstances under which that would happen.

Q. Can you give me some examples?

A. Well, let me explain where -- where I'm coming from on this.

Q. I'm asking you to answer my question, sir. And the question, I think, is a pretty simple one.

A. Would you repeat it?

Q. Sure.

MR. DOYLE: Can you repeat it, please.

(Question read)

MR. SCOTT: Objection. That's a simple question, Scott?

MR. DOYLE: Can we go up a little bit?

Q. My original question was can Informatica use more than one core at a time?

A. And I think my answer was under certain circumstances, I believe it can use more than

one core at a time.

Q. And so have you identified any of those certain circumstances in your analysis?

A. So specifically, there are two sources where I'm getting this information from. One is from the sort of hourly data that we already talked about. The second is a test that Ms. Frederiksen-Cross mentions in her initial report where she -- she runs a specific version of the Informatica software on a specific hardware configuration that does not match how Informatica was deployed on the RELX system, where there appeared to be, over a period of ten minutes, certain times during which multiple cores are used at once by the software.

So at least in -- in that specific circumstance, which I don't understand all of the characteristics of it, but there seems to be at least that circumstance specifically.

Q. So you mentioned the hourly data. What does that tell you about Informatica's use of more than one core at a time?

A. Well, it at least suggests that the software was capable of using multiple cores at the time, though I don't have any data about whether or not it used or was executing on multiple cores at a time from that data.

Q. Did you ask for that data?

A. My understanding is that it doesn't exist. I was trying to find the best data available to see records of the Informatica software executing on the ICCE platform.

Q. Do you think that data would have been helpful to your analysis?

A. Yes, I think so. The more granular the data, the better.

Q. You mentioned that your second source was the test that Ms. Frederiksen-Cross did; is that correct?

A. That's my recollection, yes.

Q. And is it your understanding from that test that Informatica uses more than one core at a time?

A. Well, the way I would say it is in the specific circumstances of that ten-minute test, it does appear that Informatica software, in that environment, was able to -- to execute on multiple cores at once.

Q. Based on that environment, do you have any reason to believe that Informatica would or would not use more than one core at a time on the RELX platform?

A. Well, there's a lot of differences between the configuration of the test that Ms. Frederiksen-Cross did and the RELX platform. Specifically, the ones that come

to mind are the number of cores on each of the machines in Ms. Frederiksen-Cross's test, as well as the type and number of data and workflows that were executed as part of that test.

Those two machines in Ms. Frederiksen-Cross's test had 16 cores on each of the machines, and there were also, I believe, 16 workflows that were configured to run using the Informatica software. So I'm not sure whether that circumstance ever occurred in the RELX environment, so I'm hesitant to say that -- that that would inform our understanding over a period of years, say, for a test of ten minutes under very specific circumstances.

Q. Well, does it inform you of whether or not the Informatica software uses more than one core at a time?

MR. SCOTT: Objection. Asked and answered.

A. Well, it helps me understand that it could in very specific circumstances, but I don't know what it would do in general, and I don't know what it would do specifically with respect to the RELX environment.

Q. So you have no idea, as you sit here today, in the RELX environment, whether the Informatica software can use more than one core at a time?

A. Well, I think it might have, but I don't know whether it did.

Q. So you're just speculating?

A. Well, I don't have any data about what actually was executed on which core, at what time, and I don't have data from a configuration that was -- that was similar in -- in certain ways to the RELX environment.

Q. Did you ask for that data?

MR. SCOTT: Objection. Asked and answered.

A. So I asked for as much data as we could get about what actually executed at what times on the CPUs and CPU cores that were on the ICCE platform, and my understanding is that I received the best data that was available. And, in fact, the initial data I received only spanned through, I think, 2016 at some point, and so I asked for the data for the last year and some change just to make sure I had complete data through the time where the Informatica software was part of the ICCE platform.

Q. Sir, I'm just, you know, a little bit confused because, you know, how can you provide opinions with respect to the benefit of the Informatica software on the ICCE platform if you don't understand how the cores were used by the Informatica software or when they were used?

MR. SCOTT: Objection.

A. So the way I approached the question of

benefit is essentially what would have happened if during the time periods where the Informatica software as part of the ICCE platform had access to more cores than the number of cores that RELX was licensed for, what would happen if during those same time periods and for that data, if instead the number of cores available to the ICCE platform were the same number that was available to RELX during -- or available to RELX through license agreements.

And so in order to answer that question, I was more concerned with overall utilization and not specifically which cores were executing the Informatica software. Because the thing that mattered to me is whether the system could -- or would have likely been able to handle that same amount of document processing over that same period of time. The details of -- of how it would have processed it weren't relevant to that question, as far as I could see.

Q. Did you look at things, for example, like processing speed, delays, things like that associated with the system?

A. So I believe I did inquire at one point about how long it took to process certain documents, but I understand that that data wasn't available for -- for the

time periods in question. And if -- if it is somewhere, I'm happy to look at it, because, you know, how long it takes to process documents does seem like another way to get insight into the question of the benefit.

Q. And in terms of processing documents, would it matter whether the Informatica software could access more than one core at a time?

A. Sorry. Would you repeat that?

Q. Sure.

MR. DOYLE: Can you read it back, please.

(Question read)

A. So because I'm comparing a situation where Informatica software had access to more cores to a situation where it had access to fewer cores, in both of those circumstances, I'm assuming Informatica -- the Informatica software as part of the ICCE platform will process documents in a specific way. I don't necessarily know the specifics of what that way is, but I can look at the degree to which overall cores were utilized to determine whether there was, you know, a meaningful difference between those two circumstances. So I don't feel as though I need to know the nitty-gritty details about which cores were utilized and which cores may not have been.

Q. Well, might those details, for example, as to whether Informatica could access one or more cores at any particular time, that -- wouldn't that be relevant to know whether there are processing delays in the system?

MR. SCOTT: Objection.

A. Well, like I said, the Informatica software in one circumstance versus the other is the same software. It's just a question of which -- I'm sorry -- in general, how many cores -- or rather, what is the kind of load on the system in order to process those -- those documents.

So, I mean, let's consider two circumstances. Supposing Informatica software only used one core of each machine, just hypothetically --

Q. Uh-huh.

A. -- then we could look at, you know, before and after -- or I'm sorry -- in the two circumstances where the Informatica software had access to more cores or fewer, what -- what the difference was between those two circumstances.

And similarly, if the Informatica software was able to use more than one core at a time, you could still look at those two circumstances and come to a conclusion about whether there was a difference in the -- the

overall utilization of -- of the -- of the servers and the CPUs.

Q. But wouldn't it matter with respect to processing time and other benefits that you might get from having more cores?

MR. SCOTT: Objection. Asked and answered.

A. Well, what I'm saying is to the extent the Informatica software was able to take advantage of multiple cores, that will be the same in both circumstances I'm describing: the one where it has access to more cores and the one where it has access to fewer. So that's the thing we're going to look at with respect to whether there was additional benefit.

Q. Well, might it change, though, if you go from like 104 CPU cores down to 56 CPU cores?

MR. SCOTT: Objection.

A. You said wouldn't it change? I'm not sure what you mean by "it."

Q. I mean "it," the processing time, for example, and whether or not you're using -- whether the Informatica software was accessing one or more cores at a time.

MR. SCOTT: Objection.

A. So the question of whether the processing time

would change, like I said, I think processing time would be interesting to look at. I don't think it's available. But it's the processing time that matters, and understanding exactly why those processing times occurred doesn't necessarily matter. What matters is what actually happened. How the Informatica software may have helped the ICCE platform to achieve those processing times, that's probably not -- or rather, that's not the first thing I would look at.

Q. So in your opinion, the thing -- the only thing to look at, or the most important thing to look at, is the utilization data you looked at?

A. Well, I would say that's the best data available to answer this question. Like I said, looking at processing times would be -- I think would be relevant, and I don't think that data exists. So in the absence of that, I looked at the best available data that I could get access to and drew my conclusions from that.

MR. SCOTT: Scott, I am sorry. I've been waiting to see a transition, but I am about to explode. I have to take a break and go to the restroom. I am so sorry. Is it okay if we go off the record for a few minutes to use the restroom?

MR. DOYLE: Sure. Just a quick break?

MR. SCOTT: Yeah, quick.

THE VIDEOGRAPHER: The time is 11:16 a.m. This is the end of Disc 1. We are off the record.

(Recess)

THE VIDEOGRAPHER: The time is 11:33 a.m. This is the beginning of Disc 2. We are on the record.

Q. Mr. Rucinski, we were talking about processing time before the break. And when you're referring to processing time, do you mean CPU or the total time to process a document?

A. I think in my answers earlier I was referring to the time it takes to -- for a document to go through the entire processing of the ICCE platform.

Q. Okay. And it's your understanding that time was not made available to you?

A. My understanding is that that data for processing time for documents is not available.

Q. Okay. With respect to processing time that you're referring to, what components make up that time?

MR. SCOTT: Objection.

A. Well, I'm defining it as the amount of time it would take for a document to go from being available from a vender, or some other source for the ICCE platform, to begin to process, all the way to that document being

available for customers of RELX to access.

Q. And what are the components of that time?

MR. SCOTT: Objection.

A. When you say the "components" of the time, do you mean the components of what would -- do you mean what would happen over the course of that time?

Q. Well, what makes up the processing time? Does I/O, for example, make up part of the processing time?

MR. SCOTT: Objection. Compound.

A. Okay. So are you asking about what components of a computer system would make processing take certain amounts of time?

Q. Sure.

A. Okay.

Q. In this instance. I mean, you know, in this case and with Informatica software.

A. Sure. So when -- when a computer accesses a file on a hard disc, there is a time between the processor kind of sending the request for, okay, I want this data at this specific location on the hard disc. There's some translation between the processor doesn't know exactly where it is on the hard disc but it wants, you know, the file associated with this particular identifier, say, and then there's a time that it takes on

a hard disc, for instance, for the hard disc head to find the appropriate platter within the hard disc to find exactly where that file is located. That file might then actually span multiple sections of that platter, and so the head might have to go back and forth in order to read it and then leave it into memory.

So yes, the input/output process, which I think is what you're referring to by I/O, would be something that would take some time for any file accessed on a hard disc, which I think the Informatica software would -- would have to do as part of its execution of ICCE platform at certain times.

Q. Is CPU another part of that time?

MR. SCOTT: Objection.

A. So I would say that it does take time for a CPU to execute its instructions, and so that would be another thing that would take time when a document is being processed through the ICCE platform using the Informatica software.

Q. And what about wait time?

A. What --

MR. SCOTT: Objection.

A. -- do you mean by "wait time"?

Q. Have you ever heard the term "wait," w-a-i-t,

used with respect to a computer system?

A. Yes.

Q. And what's your understanding of it?

A. My understanding is wait time can sometimes be used to refer to when there is a resource that is required to do something but it's unavailable for some reason. That resource could be something like -- it could be that another process is using a particular file. Maybe another process is using a different port. There could be a number of circumstances under which a resource wouldn't be immediately available and therefore a process requesting it would have to wait for some amount time.

Q. And would wait time, as you understand it, be part of the Informatica system as well?

MR. SCOTT: Objection.

A. Wait time's a little broad of a concept but, I mean, probably. At some point, there were probably some sort of bottleneck on a resource that the Informatica software might need over a period of years. But I don't have any specific evidence for specific resources that would be unavailable at certain times.

Q. What is the comparative ratio of CPU time to I/O time in the RELX ICCE processing platform?

A. Would you repeat that?

MR. DOYLE: Would you repeat that, please.

(Question read)

A. Do you have a particular time period in mind for that question?

Q. Pick any time period that you're aware of.

A. As I sit here right now, I don't know what the ratio is for CPU time to I/O processing time for the different versions of the ICCE platform and the Informatica software within it. I think that would relate to specific components of the software, but I don't have that information as I sit here right now.

Q. Did you ask for that information?

A. I don't remember asking for it, as I sit here right now.

Q. Did you look at or understand -- do you have an understanding of the comparative ratio of CPU time to wait time in the --

A. I understand --

Q. -- in the -- in the RELX ICCE processing?

A. Well, I understand what it means, but I don't know what that ratio would be for the various implementations of the ICCE platform as well as the various versions of the Informatica software.

Q. Did you ask anyone for that?

A. As I sit here right now, I don't recall asking for that.

Q. Did you ever determine how much wait time there was in the RELX ICCE platform?

A. Do you mean over -- do you mean over a particular period of time or just in general?

Q. In general.

A. I don't remember ever looking into the specific amount of wait time that -- or wait time ratios that would be present in the ICCE platform using Informatica software.

Q. What about queue time?

A. Would you define "queue time"?

Q. What's your understanding of queuing?

A. Well, a queue is a data structure in which you put items into the queue and take them out --

Q. Right.

A. -- in a first-in-first-out such that the first item that you put into the queue is going to be the first one that you remove from it.

Q. Right. And what do you understand queue time to mean as it relates to a queue?

MR. SCOTT: Objection. Asked and answered.

A. Well, I'm not -- I'm not exactly sure. It

could be something like the amount of time it takes for an item in the queue to -- to enter and then leave the queue, for instance.

Q. Sure. And that happens in the Informatica platform, right?

A. I'm not aware of a specific point in a specific component where that might happen, but documents are processed through it, so there -- there could be a queue for the documents, for instance.

Q. Sure. And did you have any information on that queue time for documents?

A. Well, we're first supposing that there is a queue. I did ask for the amount of time it takes for documents to be processed through the system, so if that's our definition of queue, then that would be related.

Q. Well, that's not my definition. I'm going with your definition of queue. Same answer?

A. Sorry. Would you repeat that? Would you repeat the question?

Q. Sure. I was just asking --

MR. DOYLE: Can we go up? Down a little bit.

Q. Yeah, my question was: "And did you have any information on the queue time for documents?"

A. So are we talking about -- well, let me first state, I'm not aware that there is a queue for documents before it enters the system, but supposing that there is, I'm not aware of any information related to that.

Q. And would queue time be related to whether or not the Informatica software could access one or more CPUs at any particular time?

MR. SCOTT: Objection.

A. Could you say that one more time?

MR. DOYLE: Can you repeat that, please.

(Question read)

A. I think it's related to whether the software could access multiple CPUs in the amount of time, but it sounds like a question about its capability.

Q. Well, would queue time vary based on whether the software could access one or more CPUs at any one time?

A. Depends how many documents are going through the queue in one.

Q. Is it a factor you looked at?

MR. SCOTT: Objection.

A. Is it a factor I'd look at for -- for what?

Q. An analysis like this analysis that you performed here.

A. Well, I would look at it if it made -- if it was -- it was material to the degree to which an actual benefit were accrued.

I mean, like I said, the first thing I'd look at is could the system handle the same documents with fewer CPU cores available. And then if it seemed like it couldn't for some reason, it might be interesting to determine why and whether that might be related to whether there was a bottleneck with respect to the CPU cores or if there was a bottleneck with respect to something like I/O time or wait time or the amount of memory that was installed or the amount of hard disc space that was available or other factors that would potentially be a bottleneck in the process.

Q. Did you ask for that information?

MR. SCOTT: Objection.

A. I did see information related to hard drives filling up. But because after looking at the utilization data and looking at whether or not there was a benefit accrued generally, it didn't make sense to me to look at these other factors that might play a role if there was -- if there was a bottleneck.

Q. So you said the -- the hard drive "filling up." Does that create a queue?

A. Well, let me clarify: So some of the e-mails that I saw related to specific directories filling up, I think many of them were in a subfolder of tmp. That's t-m-p; it's a place where an operating system puts temporary files.

So your question is would that create a queue for the hard disc. It might. If -- generally, if your hard disc is full, your computer will cease to operate very well at all. There just isn't space to hold data anymore. So unless you have another place to put it, generally, what you'd have to do is remove some of the data from the hard disc before you could proceed, at least in a functionally operable way.

Q. What about if no CPUs were available? Would there be a queue time?

A. If no CPUs were available? Well, if there was a queue that was -- that was present -- so hypothetically, if there were a queue that were present when the system was, for some reason, unable to process a document, then you might fill up a -- or you might add a document to a queue. But I'm not sure if the system would instead just not access that document from its source, or whether it would do -- or implement some other solution to it.

Q. When you provided that answer, did you provide that answer with respect to Informatica and the ICCE platform?

A. That's the example I had in mind when I was answering the question.

Q. Did you ever investigate that with respect to the Informatica and the ICCE platform?

A. Well, I didn't investigate whether in a hypothetical situation there were no CPU cores available because my understanding is that that was never the case.

Q. Earlier we talked about the comparative ratios. Do you recall that?

A. Yes.

Q. And would those comparative ratios have been helpful in this case to determine the ability to run conversions in parallel across more cores?

A. Would you my fresh -- would you refresh my memory as to what those ratios were?

Q. Sure. CPU time to I/O time, for example. CPU to wait time.

A. Let me think about that for a moment. Well, I think those ratios might be relevant to -- to why certain circumstances may have arisen in the -- in the data with respect to CPU utilization, but given that I was looking

at CPU utilization, and it seemed like the -- or actually more documents were being processed with fewer cores, it didn't really seem like it would influence my opinion in terms of why the data worked out that way. These ratios seem like something that would come up if there were a different bottleneck that would -- that would affect the degree to which the system was able to process documents.

Q. You know, I had asked you before whether CPU cores -- if you ever identified an instance where CPU cores were not available. Do you remember that?

A. I don't think you asked specifically if I recalled a time like that.

Q. Okay. Let me ask it. Do you recall a time when CPU cores were not available?

A. Well, there were times when there were certain CPU cores that were unavailable because certain servers weren't available.

Q. And so all the CPU cores on that server would be unavailable?

A. The way I'm using the word "unavailable," if the server is unavailable for processing, then yes, the CPU cores are part of the server and they would be unavailable as well.

Q. And so in that instance, a task that had been

assigned to that server would be queued until the next one became available, correct?

A. It depends how you implement the queue in this hypothetical example. If there were other servers available that do have CPU cores, I imagine such a system might send that task to that other server instead.

Q. Well, when you say "such a system," I'm talking about the Informatica software as we sit here today. Do you have any idea what the Informatica system does when a server becomes unavailable?

A. I don't know specifically what it would do in that circumstance. I imagine it would -- if it's designed well, it would probably try to take advantage of resources that it did have available.

Q. But you're speculating?

A. Yes. I didn't look specifically at how it would function in that manner.

Q. Did you ask anybody?

A. On the specific question of whether if a given server were unavailable would it then use a different server under specific circumstances?

Q. Or exactly what it would do.

A. Would you restate that question? I lost the thread.

MR. DOYLE: Can you please read the question.

(Question and answer read)

MR. SCOTT: Objection.

A. So I didn't ask whether under a very specific -- pecif -- excuse me -- a very specific circumstance what the Informatica software might do. There would be a number of different factors to consider; for instance, at what stage in the workflow is the document being processed, how far along is it.

I did see some references to, I think, in my conversations with Mr. Groff, he mentioned that there was some ability of the Informatica software to provide information about where certain documents were in the process and whether there were certain errors in certain places. I imagine one of those errors might -- or could hypothetically be related to a server going offline and not being available any longer.

Q. Did you ask RELX for that information?

A. Information about what Informatica software would do in that circumstance?

Q. Uh-huh.

A. I don't recall ever asking for that specific information.

Q. Did you identify bottlenecks in the

Informatica software on the ICCE platform?

MR. SCOTT: Objection.

A. So I looked at the CPU utilization, and that seemed to indicate that the system was able to process similar documents with fewer cores in the same manner that it processed the software with more cores. If there were a bottleneck that would otherwise prevent the system from processing documents as well, that would seem to be unrelated to the question of how many cores were at issue because I understand that's -- that's what the licenses are tied to.

For instance, if there were a different bottleneck that prevented the system from being able to utilize a certain number of cores fully, then it would perform in that same way because of that bottleneck with -- with fewer cores.

Q. You mentioned it's your understanding that's what the licenses related to. What do you mean by that?

A. Well, my understanding is that the license that RELX obtained from Informatica, for instance, counted -- and I believe the specific language was CPU cores that the -- the license provided for the software to be installed on computers that had that certain number of CPU cores as in the agreement.

Q. And were those limits -- the number of CPU cores limits on the scope of the license?

MR. SCOTT: Objection. Calls for a legal conclusion.

A. I do think that calls for a legal conclusion.

Q. I'm not asking you as a lawyer. Are those limits? Meaning you can't go over that amount that specify the number of CPU cores?

MR. SCOTT: Same objection.

A. Well, it's my understanding that the license allowed for a certain number of cores to be available for the software, and -- and it didn't programatically prevent someone from installing software on servers that had a greater number of CPU cores available.

Q. I didn't ask whether programmatically. I asked do you understand that to be a limit of the license in terms of how many CPU cores RELX could use.

MR. SCOTT: Objection.

A. So it's my understanding that if -- if RELX had -- if the Informatica software were installed in such a way that it had access to more cores than were specified by the license, that would -- I understand would exceed that -- that number as provided by the license.

Q. Do you have an understanding of the different workload types LexisNexis processed with the Informatica software?

A. Sorry. Did you say "workload" types?

Q. Yes.

A. I understand the Informatica software had workflows that were processed. I guess I don't know what you mean by "workload" types.

Q. Yeah, the workflow types. Do you have an understanding of the different workflow types LexisNexis processed with the Informatica software?

A. I did talk to Mr. Groff about different workflows that -- that were part of the overall process in the ICCE platform.

Q. Other than Mr. Groff, did you receive any information on the different workflows?

A. There was a document that -- that I received that sort of summarized the -- the different workflows that -- that were part of the process.

Q. What's that document?

A. There are excerpts in it in my report that shows the different implementations for certain types of workflows. Some were data transformation for Informatica, like we talked about, and others were Java

and Perl implementations from RELX.

Q. What content was carried by those workflows?

A. I think in general they were all related to processing individual documents that would, at the end of the process, be made available to RELX customers on either Lexis.com or Lexis Advance.

Q. But I'm talking about the specific type of content. What type of content?

A. Are you asking about the specific types of documents that are processed?

Q. Yeah.

A. So I understand that -- I understand that there were certain types of documents that were processed by the -- excuse me. I understand that there were certain types of documents that were processed by the ICCE platform and certain other types that weren't, at least with respect to their content. So it's my recollection that mostly the documents that were processed had to do with legal cases in some way, and that there were other documents that were eventually made available on Lexis.com and Lexis Advance that didn't go through the ICCE platform and -- and were outside the scope of that category of document.

Q. What types of legal content went through the

Informatica portion of the ICCE platform?

A. Well, I'm not sure if all of it did, but some examples of content that would have gone through, I believe, were things like documents related to different legal cases that were in progress, different statutes that -- that were made available. I believe there were others, but those are the ones I recall as I sit here right now.

Q. Is it your understanding that's generally public information?

MR. SCOTT: Objection.

A. My understanding is that at least statutes are generally public information. I think sometimes you may need to pay a small fee to access certain other court documents. At least that's been my experience.

Q. So going back to workflows, what were the primary workflows?

A. Let me see. So I discussed this in some detail in my report. As I sit here right now, my recollection is they were workflows related to acquiring the document from the vendor that -- or whatever source the document came from. Then there were workflows related to extracting the text from those documents and formatting it in a way that was consistent with other

documents that were processed by the system. And then there were also workflows related to storing that textual information in a manner that was, again, consistent with other documents that were processed and so could therefore be made available to customers of RELX.

Q. And what was the desired processing time for each of those workloads -- or workflows?

MR. SCOTT: Objection.

A. I'm not sure what the specific objective or ideal processing time for specific workflows were. My conversation, I think in that e-mail in my report from Mr. Groff, he mentioned that times around three hours were kind of less than what they -- what they cared about in terms of processing time.

Q. So for each of those workflows, your understanding is that the desired processing time was less than three hours?

MR. SCOTT: Objection.

A. That's not my understanding. I think -- I don't know what the exact processing time for individual workflows would have been.

Q. Did he ask for that?

A. Did Mr. Groff ask for that?

Q. No. Did you ask anybody for that information?

A. I did not ask about the specific processing times for the different workflows.

Q. So for the different workflows, how long did it take to process the workflows before the cores were reduced from 104 to 56?

A. You're asking for a specific time period or dates or --

Q. Answer any way you can. I mean, I'm just asking you how long did it take to process each of the workflows, you know, at full 104 cores and then after when you went to 56 CPU cores?

MR. SCOTT: Objection.

A. Well, it would probably depend on how you measured it. I mean, I don't know if -- if a workflow has an error, would you discard that or not. My understanding is that over time, the platform changed quite a bit to use less Informatica software than it did at the beginning, but I haven't seen any data with respect to exactly how much time it took for documents to go through the ICCE platform either in general or through specific workflows.

Q. When you say the exact time, did you see some documents or any other information received that provided how long it took to process each of the workflows?

A. There may have been some discussion I saw in passing, but I don't recall it, as I sit here right now.

Q. Did you do a comparison between the time it took to process the workflows both before the core reduction and after the core reduction?

A. Well, I asked for information on how long it took to process documents through the system, and the workflows are part of that system. So I don't remember seeing any information related to how long it took to get through the system, and so that wasn't part of my analysis.

Q. So you say that Mr. Groff told you that -- well, that let me -- let me back up. You mentioned three hours. What was three hours related to?

A. Well, I was asking, I think -- and this would be clearer in the actual e-mail -- but I was asking with respect to the granularity of the utilization data that we had on a per server basis and whether or not it would matter if there were, you know, a small spike in utilization within that hour because we only had -- because we only had the average.

So my understanding is that in response to the question of, well, did it -- does it matter that we only have granularity of -- of an hour for these CPU

utilizations, Mr. Groff's response was, well, we only -- we only really care about granularities of three hours, and so granularities of one hour is -- is enough to look at that data.

Q. What does it mean a granularity of three hours?

A. So the data that was produced, when I say it has a granularity of one hour, what I mean is that it provided the average CPU utilization for a server coveraged over across one hour, you could imagine different time periods over which that might be averaged such as two hours or three hours; you could imagine granularities that are smaller than that; you could imagine an instantaneous, you know, measure if you wanted, although that would -- I'm not sure how useful instantaneous would be.

Q. And Mr. Groff said that he was only concerned about average CP -- or that RELX was only concerned about average CPU over three hours?

A. That's my recollection. If I look at the -- the e-mail, I'll be able to answer that better.

Q. Did you ask if it was ever over three hours difference between 104 cores and 56 core implementation?

MR. SCOTT: Objection.

A. I'm not sure what you mean by "if it was over three hours."
Q. Meaning did it take longer than three hours.
MR. SCOTT: Objection.
A. I'm not sure what you mean by "it."
Q. Did he look at four hours or five hours?
MR. SCOTT: Objection.
A. For what? I'm not sure what the context is.
Q. The utilization.
A. Well, so if you have granularity of the data on an hour-by-hour basis averaged over those hours, you could then take three of those instances and take the average, and now you have the average for three hours.
Q. Okay. Did you ask Mr. Groff if the big peak or the big spike would matter in the operation on the system?
A. Which big spike are you talking about?
Q. The big one you referred to in your answer.
A. I referred to hypothetical spikes.
Q. Oh, you never saw a pike in any of the data you looked at?
A. Is that a question?
Q. Yeah.
A. Well, so spike is referring to -- or

hypothetical spikes I was referring to in my previous answer were on the order of a time period less than an hour. What I'm saying is that over the course of an hour, because I only had the average, I couldn't tell whether there were periods of time over the course of that hour where a CPU -- or rather -- let me be precise -- whether individual CPUs or cores on the server would have been utilized by more or less than an average, which could be something like zero for certain periods of time, or it could be more like 100 at certain periods of time.

Q. Oh, so the utilization data you looked at doesn't tell you anything about spikes?

A. I wouldn't say it doesn't tell me anything, but it doesn't identify how much utilization was -- let me rephrase that.

It didn't show how -- how CPUs or CPU cores were being utilized for periods of time that were less than the hour-by-hour average that we do have.

Q. So it doesn't tell you how it reacted when there was a big spike?

MR. SCOTT: Objection.

A. Well, it tells me over the course of an hour.

Q. Yeah, but not for a short period of time. If

the spike was over a shorter period of time, you have no information on that when you're just looking at utilization of an hour?

A. I wouldn't say --

MR. SCOTT: Objection.

A. I wouldn't say I have no information about it. That spike that we're hypothetically talking about on a period of time less than an hour is reflected in the average for that hour.

Q. Oh, sure. But you don't know whether it actually occurred or not and when it occurred, correct?

A. I could probably put limits on it, but I wouldn't know the exact details of when it occurred within that hour.

Q. When you say you could "probably put limits on it," wouldn't that be speculating?

A. No. Let me explain. What I mean by that is if you have the average CPU utilization for a server over an hour at, say, 50 percent, then it follows that you couldn't have all of the CPU or CPU cores of that server utilized for 100 percent utilization for more than half of that time. For instance, if you had the CPU or CPU cores utilized for 100 percent of the time for a half hour and then zero percent for the rest, then you'd have

your 50 percent average. But as soon as you have them hypothetically utilized for 100 percent for more than 30 minutes, your average would be over 50 percent which would be inconsistent with the average hourly data that we do have.

Q. So did you actually analyze any of the big spikes that occurred?

MR. SCOTT: Objection.

A. Well, the question I was trying to answer was whether -- or to what degree RELX benefited from having the Informatica software deployed on servers that have access to -- to cores that were a greater number than the number of cores that they were licensed for.

So the spikes of -- that hypothetically could have existed over periods of time for less than an hour didn't matter to the analysis, as Mr. Groff said in his e-mail, with respect to CPU utilization, certainly anything less than three hours wouldn't -- wouldn't really affect anything, and so it didn't affect my analysis. And also, the data for granularities less than one hour doesn't -- doesn't exist, as far as I know.

Q. Did you ever ask if obtaining more granular information about big spikes would matter to RELX?

A. Well, I think that question was incapsulated

in my question to Mr. Groff about the time periods over which CPU utilization matter. I'm happy to look at the e-mail if it would help to get more clarity on that question.

Q. Did he provide you with any information about spikes or when they might occur?

A. Well, we're still talking about spikes on the order of less than an hour. As far as I know, there was no data on a granularity less than one hour, and I don't have any data related to that.

Q. Okay. If the CPU average was 100 percent for an entire hour, would that mean that all cores were busy in that hour?

A. If it was exactly 100 percent for an hour? I can't think of a circumstance under which you would have a full 100 percent utilization for a server that had multiple cores where you didn't have some of those cores -- or rather, you didn't have all of the cores active for some of that time at least.

Q. That's not the question I asked you.

A. Would you repeat the question?

Q. If CPU average was 100 percent for an entire hour, would that mean all cores were busy in that hour?

MR. SCOTT: Objection.

A. When you say "busy," do you mean utilized?
Q. Busy.
MR. SCOTT: Objection.
A. What do you mean by "busy"?
Q. Unable to process more data.
A. Well, there are certain instances where CPUs can report utilization higher than 100 percent. It doesn't happen often but it does happen sometimes. So if the question is if CPUs on average were reporting 100 percent utilization over the course of an hour, could they have processed more data, I think the answer might be yes. It depends on the circumstances.
Q. If it's over 100 percent, does that mean all the cores used, and each core used, is at 100 percent?
MR. SCOTT: Objection.
A. Do you mean that they were at 100 percent for the entirety of the hour?
Q. If you -- you get a thing that says over the hour it's 100 percent, right, for a server. And I'm just asking you does that mean all the cores used in that server were, for each of those cores, was it at 100 percent?
MR. SCOTT: Objection.

A. Well, like I said, sometimes CPU cores can be slightly over 100 percent for periods of time, so it could be the case that maybe half of the cores were at 99 percent and the other half were at 101 percent which would average to 100 for the hour.

Q. But it's also the case they could be 100 percent, right?

A. It could be the case that if your average shows that all the CPU cores on average for an hour were 100 percent, it is possible that all of them were at exactly 100 percent for the entirety of that hour.

Q. Do you know what a desired processing window is?

A. I have the definition in my head of what that could mean.

Q. And what is that?

A. Well, it sounds like an amount of time during which you would like to process something.

Q. And did you have or receive any information about the desired processing window or service level for each of the workload types?

A. Well, like I said earlier, I did ask for information about how long it would take to process a document through the system, and I understand that that

information wasn't available, or at least I didn't see it.

Q. Who did you ask?

A. I believe I asked -- it was either Mr. Groff, or I think Julie, who is sitting right here, or both.

Q. And you didn't receive any information?

A. No, I didn't receive any information about how long it would take to -- or how long it did take historically to process documents through the system.

MR. DOYLE: Okay. Could we get the expert report.

MR. SCOTT: It might be a record, Scott. It's 12:17. We're marking our first exhibit.

MR. DOYLE: Yeah. And this is going to be Rucinski 1, and it is a copy of Mr. Rucinski's original expert report from May of this year, 2018.

(Exhibit No. 1, Rucinski expert report marked for identification)

MR. SCOTT: This one has no highlights.

MR. DOYLE: What's that?

MR. SCOTT: This one has no highlights.

MR. DOYLE: Okay. Thank you.

Q. Can I direct your attention to Exhibit 6 -- I'm sorry -- Exhibit H, please.

A. Okay. I'm there.
Q. Okay. What do you understand this to be?
A. This appears to be Exhibit H to my expert report.
Q. And what is it?
A. It is an e-mail from Julie to me that documents responses that Mr. Groff provided to questions that I provided him through counsel.
Q. Okay. Can you turn to page 5, please.
A. Okay.
Q. If you go to what's been -- what's 5, you see where it says "Questions not based on documents already produced"?
A. Yes.
Q. Can we go to No. 2.
A. Okay.
Q. Do you see where you asked -- I believe that's your question, right? -- where you say: "What was the general time frame that LexisNexis wanted documents processed through Informatica's B2B Data Exchange software to be completed?"
A. Yes, that is the question I posed in this e-mail.
Q. And what was the answer provided?

A. The answer provided was from Mr. Groff: "The expected time frame varied across the content streams as well as the number of files in a given batch. The performance and reliability of the ICCE Informatica grid was very poor when Informatica originally implemented the ICCE 1.0 platform. The Informatica grid would crash frequently, and some of the DT services would run for days or sometimes corrupt, in quotes. DT services would corrupt the entire DT service to fail and stop all DT processing. Due to these issues, LexisNexis was unable to determine valid expectations for content processing in the ICCE 1.0 platform.

"With a lot of re-engineering of the Informatica PowerCenter maps and DT service calls by the CCP RELX team, we were able to get the platform considerably more stable and increased overall performance with the ICCE 2.0 release. ICCE 2.0 was implemented in production on June 2014, but only some of the content streams were migrated from ICCE 1.0 to ICCE 2.0 due to time and budget considerations."

Q. So looking up at the top, do you see where it says "the expected time frame varied across the content streams as well as number of files in a given batch"?

A. I do see that.

Q. Don't you presume from that answer he has some information that relates to time frames?

A. Well, he answered my question with more than just that number -- or rather -- sorry -- more than just that sentence, so given that I asked the question and this was his response, I assume this is what he had to provide for an answer.

Q. But he knows that the expected time frames varied across the content streams, right?

MR. SCOTT: Objection. Calls for speculation.

A. Well, I don't know what he knew. I know what he wrote back to me in response to this question.

Q. Did you follow up and seek the information that he was referring to: that the expected time frames varied across against the content streams?

A. Well, I took his answer to the question posed to be a complete one here, and he didn't provide what those expectations may have been.

Q. Did you ask about the differences between content streams in terms of expected time frames?

A. I didn't follow up on this question.

Q. Does it appear that RELX had some information as it related to expected time frames for content streams --

MR. SCOTT: Objection.

Q. -- that was available to RELX?

MR. SCOTT: Objection. Calls for speculation.

A. I don't know what information RELX had.

Q. Well, did you ask him?

A. Well, I asked Mr. Groff in his e-mail.

Q. Yeah. But did you follow up and ask him about expected time frames?

MR. SCOTT: Objection. Asked and answered.

A. I already said I didn't follow up on this specific question because I had an answer on it already.

Q. Would that have been helpful to get some information on the expected time frames in terms of forming part of your opinion?

A. Well, the thing that I wanted to see was records of how long it took for files to get through the system. So expected time frames are somewhat interesting, I suppose, but to the question of how -- how might have the time it took for documents to actually make it through the system, that's not a question about expectations. That's a question about what actually happened historically.

Q. Sure. But wouldn't it be -- wouldn't it be helpful to know the actual expected time frames from RELX

for these various content streams so you could compare how long it actually took to get across the system and compare it to the expected time frames?

A. Well, if I don't know what the actual data was for how long the files took, I don't have anything to compare it against.

Q. Okay. So what you're saying is because you had no information about how long it took for the data to actually get through the system, you didn't need to know any other information about expected time frames?

A. Yeah. I think I would say because I didn't have the information about how long it actually took for documents to get through the system, I did not have anything to compare it to with respect to expectations.

Q. And you never got any records or documents about how long it took to actually get through the system, right?

A. My recollection is I did not get documents that showed how long it took documents to get through the system.

Q. Okay. Are you aware if there was any evidence that expected time frames were not met after the reduction in cores from 104 CPU cores to 56 CPU cores?

A. I don't remember seeing anything like that. What comes to mind right now is just what's in the e-mail here and Mr. Groff's response. But I think that was related to processes crashing with the ICCE 1.0 platform, which, in general, I think was earlier than the time period at which the number of cores in Informatica's system was reduced.

Q. So you had no time for processing document types for the system when it had 104 CPU cores?

A. I don't remember seeing documents for the amount of time it took to process documents through the ICCE platform for any period of time.

Q. Did you have any information that disclosed processing times through the platform when it was 56 CPU cores?

A. Well, like I said, I don't recall seeing any documents that -- that showed how long it took documents to get through the system for either of the -- either the time before or after the reduction in the number of CPU cores that were part of the ICCE platform.

Q. Did you ask for that information?

A. Again, I asked, in general, for -- if there's information or documents about how long it actually took documents to get through the system, then that would be

information I would like to have in forming these opinions.

Q. And what was the response?

A. I never got any documents that showed that, and my understanding is that those records don't exist.

Q. And who told you those records don't exist?

A. I don't remember exactly. I remember asking for it and never getting it, whereas I asked for increased -- or I'm sorry -- the utilization data for different time periods and did get that information, and so I don't remember exactly how it transpired, but my assumption was that because I asked for it and didn't get it, it must not be there.

Q. Who did you ask? Counsel?

A. I believe I -- I believe I asked counsel, and I think I asked Mr. Groff as well at some point.

Q. Do you recall there was a period in time when the parties signed an amendment that retired 16 CPU cores?

A. My recollection is that there was a time when the CPU core license, whereas before it covered 72 cores, after the new agreement, it covered 56, if that's the one you're talking about.

Q. Okay. And what was the reason why RELX

retired those 16 CPU core licenses?

MR. SCOTT: Objection. Calls for speculation.

A. I don't know why RELX would have made that decision.

Q. Did you ask anyone?

A. No, I don't think I did.

Q. Is it relevant to your analysis and your opinions?

A. Well, the question of why parties acted the way they did seems outside the scope of computer science opinions in general, but also, I was looking at specifically what evidence there was for a benefit that was accrued, if any, to RELX when they had more cores versus fewer cores on the system, and for that purpose, I looked at the data for utilization that was available.

Q. But doesn't the actions of RELX -- the actual factual actions of RELX have any bearing on that when they actually reduced or they signed an agreement to reduce the number of CPU cores?

MR. SCOTT: Objection.

A. Well, I don't know why they did that.

Q. Did you ask?

MR. SCOTT: Objection.

A. I didn't ask them why they -- why RELX did that.

Q. Well, did you ask him why they didn't actually reduce the number of CPU cores after signing that agreement?

MR. SCOTT: Objection.

A. I don't recall asking RELX why -- why they made certain decisions because I was more focused on the actual data of what transpired and not the reasons for -- not the reasons for it happening the way it did.

Q. So it doesn't matter to you that somebody signs an agreement to actually get rid of a number of CPU cores and then they actually continue using the CPU cores? That has no relevance as to whether or not they needed to continue using the full amount of CPU cores?

MR. SCOTT: Objection.

A. I wasn't looking to the question of need. And it's relevant insofar as it is -- what -- insofar as those agreements define the amount of licensed cores that were available at certain times.

Q. Why did they stay at 72 CPU cores if they had signed an agreement that they would go down to 56 CPU cores?

MR. SCOTT: Objection. Calls for speculation.

A. I don't know why RELX would or would not do the things that they did or contemplated doing.

Q. Well, isn't it your opinion they only needed 56 or fewer cores?

A. Well, in order to process the similar data over the different time spans for which they had different amounts of cores, you know, my opinion is that they could have processed the same data just as well with -- with the smaller of the two amounts of cores available to them.

Q. Sure. So did you ask him why they stayed at 72 CPU cores and didn't drop down to 56 CPU cores?

MR. SCOTT: Objection. Asked and answered.

A. I didn't ask RELX why they made the decisions that RELX made.

Q. Why not?

A. It didn't seem relevant to the opinions because I was focused on what actually transpired and not the reasons why certain decisions were made.

Q. Couldn't there have been performance issues? They wanted to stay at 72 CPU cores because they needed to because of performance issues?

MR. SCOTT: Objection.

A. Many things are possible. I don't know why

they decided to do the things that they did.

Q. Shouldn't you have asked them?

MR. SCOTT: Objection.

A. Like I said, I was looking at whether they actually accrued a benefit, and so to answer that question, I looked at the actual data as it existed. It could, for instance, have been the case that they thought certain things that were untrue, and so that's why I looked at the data.

Q. Did you look at processing times between 72 CPU cores and a 56 CPU core system?

A. Well, I did ask, for instance, if there were data related to how long it took documents to process through the system, but --

Q. You didn't get it, did you?

A. I didn't see any documents that had that level of data in them.

Q. So it's possible they remained at 72 CPU cores because they needed to due to certain performance issues that you didn't investigate, correct?

MR. SCOTT: Objection.

A. I don't know why RELX did the things that they decided to do, so --

Q. But it's possible they could have, right?

MR. SCOTT: Objection.

Q. Due to performance issues, correct?

A. It's possible they could have acted in a number of different ways for a number of different reasons. Other --

Q. But they could have due -- they could have done it due to performance issues, correct?

MR. SCOTT: Objection.

A. They could have done it for a lot of reasons, and that is a reason, so...

Q. Do you know if the failure to actually continue to -- I mean, do you know if the failure to reduce the number of cores by 16 related in any way to critical workflows?

MR. SCOTT: Objection.

A. Are we still talking about the change from 72 cores to 56 cores?

Q. Yes, we are.

A. Would you repeat the question?

Q. Yeah. Was the -- do you know if the failure to reduce the cores was related in any way to critical workflows?

A. Well, I didn't ask RELX why they -- why the number of cores that they had deployed -- or -- I'm

sorry. I didn't ask RELX why they chose to act in ways or not act in certain ways with respect to the number of CPU cores as part of the system, so I don't know whether that would have had an effect.

Q. So you don't know whether it was -- the failure was related in any way to critical workflows, right?

MR. SCOTT: Objection.

A. Well, I don't know if I'd kind of characterize it as a failure, but, again, I was looking more at the actual data of the -- that was in the utilization for its servers, and I wanted to look at that as opposed to impressions that different companies or parties of companies had because the data is the actual record of what transpired.

Q. You've spoken with a number of members of the RELX IT staff; is that correct?

A. I've spoken with two of them.

Q. And who are they?

A. One is Dwight Groff.

Q. Okay.

A. And the other is Jeff Hoffman, I believe his name is.

Q. Anybody else?

A. Those are the only two employees of RELX that I spoke with.

Q. Do you know if Dright -- Dwight Groff and Jeff Hoffman were at RELX at the time of the amendment that reduced the licensed CPU cores from 72 to 56?

A. Sorry. I didn't hear that. Would you say that one more time?

MR. DOYLE: Can you read it back, please.

(Question read)

A. I'm not sure if they were. The reason I spoke to Mr. Groff in particular was because he was provided as the person from RELX who was most knowledgeable about the sort of technical implementation of the ICCE platform and the time periods during which the Informatica software was implemented as part of it.

Q. Do you think they're competent in their knowledge of the RELX system?

MR. SCOTT: Objection.

A. Well, I'm not here to offer opinions about competency, but they were provided as the best available employees from RELX to answer technical questions.

Q. And do you believe they did a good job answering your technical questions?

A. Overall, yeah. I mean, they did answer the

questions that I asked them, and I found no reason to doubt their responses.

Q. Do you think Groff and Hoffman knew why RELX failed to reduce the number of cores after the amendment --

MR. SCOTT: Objection.

Q. -- that required the reduction of 16 CPU cores?

MR. SCOTT: Objection.

A. Well, I don't know what -- what they know. Only they would be able to tell you what they know.

Q. Based on your experience and expertise, what is the most likely reason RELX did not reduce from 72 CPU cores to 56 CPU cores?

MR. SCOTT: Objection. Calls for speculation. Assumes facts not in evidence.

A. Well, there are a number of reasons that they might have done that, and as far as my computer science expertise, I mean, I think a lot of those reasons would fall outside the scope of that expertise, so I'm not sure why they would -- why they would have made certain choices.

Q. But, again, you didn't ask anybody?

A. Well, I didn't ask why RELX made certain

choices because I was focused on the data of what actually transpired and not the reasons for certain choices being made.

Q. Did you ask counsel?

A. Counsel is a member of the group of people that I might ask this question, and I didn't ask the question in general.

Q. Do you know if LexisNexis still met all of its service targets and performance targets after the number of cores were reduced to 56?

A. I don't know what service targets they were -- were measuring against for -- for those time periods.

Q. You didn't ask?

MR. SCOTT: Objection.

A. As far as I knew, the best data that was available was the actual CPU utilization data, and so that's the data that I focused on.

Q. But you didn't ask them if they had service targets?

A. Well, again, I asked for how long it took documents to process through the system, and because I didn't have the information, I had nothing to compare it against. That probably would have been the next question in that case.

Q. Earlier Groff had stated that the expected time frame for processing varied across the content streams. Do you remember that?

A. He said the expected time frame varied across the content streams as well as the number of files in a given batch.

Q. Okay. So given that, do you know if RELX was still able to keep the expectation times for the workflows -- I'm sorry. Do you know if the actual times for processing were underneath the expectation values, or were the expectation values after the reduction on the 16 cores?

A. Well, I didn't see data for how long it took for the documents to actually get through the system, and so I had nothing to compare expectations against, and so I didn't perform the comparison.

Q. Overall, wouldn't it have been important to know that in relation to coming to a conclusion that the Informatica software provided no benefit?

A. When you say "that," what do you mean?

Q. Knowing processing times.

A. Okay. Could you repeat it all at once? Sorry.

Q. Sure. Wouldn't it be important to know

processing times and how they compared to expected processing times before coming to a conclusion that there was no benefit due to the Informatica software?

A. Well, I would want to know how long it would take documents to get through the entire system. After I -- which I didn't see. After getting that, it might be useful to compare against either previous -- previous implementations that -- for which there was data for the amount of time it took documents to get through the system.

The expectations, I think, would be relevant. It would probably also depend, though, on how the expectations were formed. You might -- there could be a circumstance, for instance, where an expectation was -- was maybe too aggressive and maybe the company didn't actually benefit from getting to that expectation or not. But that would be something I would have to look at in more detail in an actual situation.

Q. Sir, did you have a conversation with Dr. Vellturo in this case?

A. I did.

Q. And Dr. Vellturo has been hired as a damages expert in this case?

A. That is my understanding as well.

Q. Did you tell Vellturo that the Informatica software provided no value?

A. I think what we discussed was the degree to which RELX benefited from having the ICCE platform have access to more cores than the number of licensed cores that were available to them over certain time periods. My opinions are not related to the general benefit that RELX may have accrued from Informatica -- from the Informatica software. It was directed specifically towards any difference that may have happened with respect to a different number of cores being available to the ICCE platform.

Q. Did you provide him any information relating to the benefits or disadvantages of Informatica software in general?

A. That is not something I had planned about in my report, so I don't think I provided that to Dr. Vellturo in the phone call.

Q. So what exactly did you tell Vellturo?

A. One thing we talked about was, I think, the different options that RELX had for how to implement its ICCE platform and some of the different avenues they could have taken in terms of performing that implementation.

So one avenue that I previously discussed with Mr. Groff was using a, you know, expansion of the Java and Perl scripts that they had previously been using in the new system and that Informatica was a third-party solution for at least part of the ICCE platform. And I understand they considered other options, but those were the ones that we talked about.

Q. But you told Dr. Vellturo Informatica provided no value when going from a hundred -- I'm sorry. Strike that.

Did you tell Dr. Vellturo there was no value or benefit when RELX was using 104 CPU cores as opposed to 56 CPU cores?

A. Well, generally, my opinion is that there was no benefit accrued with respect to more cores being available to the Informatica program as part of the ICCE platform over the certain time periods for which there were more cores as part of the system versus the number of cores that were available because of the license agreement.

Q. And what was the number of cores that you thought that you could adequately run the system with?

A. I don't think I ever came to a determination about what that specific number would be. I was

concerned only with whether there was a difference in the benefit between the number of cores that were available through a license agreement and the number of cores that were actually on the ICCE platform that the Informatica software, as part of that ICCE platform, had access to.

Q. Did you tell him that the ICCE platform using Informatica only required 56 or fewer cores?

A. I don't remember the exact conversation, but probably the way I phrased it is that there was no benefit in having more than 56 cores on the system because, as I explain in my report, there was actually a time during which 56 cores were both part of the ICCE platform as well as number of cores that were licensed, and there were more files processed through the system over that time period than over the previous time periods where there were more cores available to the Informatica software as part of the ICCE platform.

Q. So how did you come to that conclusion if you had no information relating to processing times or expected workflow times?

A. Well, as I explained in my report, I looked at the utilization data for the servers that were part of the ICCE platform over various time periods. I also looked at the number of files that were processed through

the system on a daily basis.

Q. And you think that that's sufficient to come to that conclusion?

A. I do.

Q. So processing times aren't important? Is that what you're saying?

MR. SCOTT: Objection.

A. I wouldn't say they're unimportant, but I don't think they're necessary to come to the conclusion that I did.

Q. In Footnote 31, you say -- of your report -- it's on page 32 -- you state: "Across all seven ICCE servers, more than 99 percent of hours have data reported over the approximately three-year time frame covered by the ICCE server utilization report."

Do you see that?

A. I see Footnote 31 which reads: "Across all seven ICCE servers, more than 99 percent of hours have data reported over the approximately three-year time frame covered by the ICCE server utilization report."

Q. What does that mean?

A. Well, give me a moment. So there were certain hours that had no data reported in the -- in the data that I was looking at. I understand that some of that

was related to servers actually not being online during those times.

Q. Okay. So it just means the data that was available to you covered 99 percent of the time?

A. I think it was somewhat more than 99 percent, but I said 99 percent to be safe.

Q. And how did you obtain that information?

A. Well, I manually reviewed the data that was available in the spreadsheet and took note of any gaps. And there were at least two types of gaps identified: one where there was an actual entry for the hour but then no data reported; and there were a handful of instances where there were hours that just weren't in the spreadsheet, there wasn't a row for them, and so I made sure to include those days -- I'm sorry -- those hours with no data for them.

So in summary, I looked through the spreadsheet. I also did some programatic verification just to make sure that there were the correct number of rows versus the time frame that was -- that was in the data.

Q. Sir, did you author your entire report?

A. Give me a moment. Don't worry. I won't flip through the whole thing.

I wrote it all except there were two sentences suggested by counsel which, in this report, are at the end of paragraph 2. They read: "This report contains confidential information including information contained in documents identified by Bates numbers that were produced by the parties to this litigation during discovery. This report is subject to the protective order agreed to by Informatica and RELX on May 26, 2017."

Counsel suggested that I include these two sentences to help protect the confidentiality of this report.

Q. So what sections?

MR. SCOTT: Objection. Asked and answer.

Q. I think you referred to two sections suggested by counsel.

A. In my previous answer, I was referring to the two final sentences at the end of paragraph 2 in this report.

Q. Paragraph 2?

A. Yes. It's on page 1.

Q. And what did counsel provide to you?

A. They provided the suggested language which I modified slightly. The language, again, is those two final sentences beginning with "this report contains

confidential information," and ending with "agreed to by Informatica and RELX on May 26, 2017."

Q. Anything else that counsel provided to you?

A. I wrote the rest of my report myself.

Q. Okay. Do you have a bunch of opinions in here with respect to your analysis about copyright registration; is that correct?

MR. SCOTT: Objection.

A. I do have opinions that relate to copyright registration.

Q. What page is that?

A. Give me a moment.

Could we also break in the next 10, 15 minutes?

Q. Sure. We'll break for lunch.

A. All right.

So I think those opinions begin around page 42, Section Roman numeral IX.

Q. And that section goes through, let's see, page 46?

A. I believe it goes through page 48. The next top level section is 10, so that's on page 48.

Q. Okay. Did counsel write you to ask you to write this section?

A. Counsel asked me to offer opinions about what -- well, among other things, what type of -- or to characterize the software that was included in the Informatica copyright registrations.

Q. Are you a copyright law expert?

MR. SCOTT: Objection.

A. Well, the opinions I'm offering are not in the area of law. They're in the area of computer science.

Q. Well, we'll get into your opinions in a minute.

MR. DOYLE: Why don't we go ahead and take our break now, and we'll come back. About an hour, since we're on a tight schedule? So it's one o'clock now. Two o'clock?

MR. SCOTT: Okay. Do you want to fill me in on if there's something on the backside? Does somebody have a flight or something?

MR. DOYLE: There are flights at 7, yeah.

THE VIDEOGRAPHER: The time is 12:55 p.m. This is the end of Disc 2. We are off the record.

(Recess)

THE VIDEOGRAPHER: The time is 2:06 p.m. This is the beginning of Disc 3. We are on the record.

MR. DOYLE: I'd like to ask the court reporter

to mark as Rucinski Exhibit No. 2 the presentation called "Unlock the Potential of External and Hierarchical Data."

(Exhibit No. 2, "Unlock the Potential of External and Hierarchical Data" marked for identification)

Q. Have you seen this document before?

A. Let me flip through it for a moment. Well, I'm not sure if I've seen this particular document before, but it does look familiar.

Q. Do you know who Brian Wisvari and Xinwei Li are?

A. I believe they're two employees of RELX. Go ahead.

Q. And do you know what their roles were with respect to the ICCE platform?

A. To sit here right now, I don't recall specifically how that you were involved with ICCE.

Q. Did you have any conversations with them?

A. I didn't have conversations with either Brian Wisvari or Xinwei Li.

Q. Did you read their -- Brian Wisvari's deposition transcript?

A. I did read through it quickly, yes.

Q. Okay. Do you see that this is a presentation

provide -- given by Brian Wisvari and Xinwei Li?

A. Well, their names are on the title slide. As I sit here right now, I don't know whether they presented it, but their names are here.

Q. Okay. Do you see the date June 5, 2013?

A. That date is on the first page of this document, yes.

Q. Okay. Unfortunately, the pages aren't numbered, so could you go to the top of the page. It says "Products to Solve Problems."

A. Okay. I think I'm on that page, yes.

Q. Okay. Do you see the Informatica B2B Data Exchange?

A. Yes. The first bullet reads "Informatica B2B Data Exchange 9.5.0."

Q. Sure. And do you understand the Informatica B2B Data Exchange to comprise PowerCenter, data exchange, and data transformation?

A. That sounds correct, it's my understanding. That's also what's on this document.

Q. What's PowerCenter?

A. So my understanding is that PowerCenter -- and I think I go through this in more detail in my reports -- but my understanding is that PowerCenter is kind of the

main part of the Informatica software where you can see what's going on in various other parts and kind of -- I don't know, the center of the software from a monitoring perspective.

Q. So it monitors. What else does it do?

A. As I sit here right now, my recollection is that it -- it's kind of the -- it's the part of the software that organizes or -- and keeps track of the other portions of it, so data exchange and data transformation being the two other parts, for instance.

Q. Does it control the data exchange and data transformation?

A. I'm not sure whether it controls it, but it is -- my understanding is it's sort of a higher -- it's at a higher level in the software than those other two components, so kind of overseeing the other two.

Q. What's data exchange?

A. My understanding is that data exchange is the component of the Informatica software that is related to actually acquiring documents for processing, and so it will be the component of the software that monitors for when certain documents are available for processing and plays a role in orchestrating how those documents enter the system.

Q. Okay. And what is data transformation?

A. My understanding is that data transformation is, I would say, at the lowest level in the software hierarchy of these three. It is the portion of the Informatica software that is responsible for actually performing, well, as its name suggests, data transformation from, for instance, the original document to a text form that is more suitable for ingestion into the larger, in this case, ICCE platform, for instance.

Q. And what is the text form that you're referring to?

A. I think in general the way the ICCE platform worked with Informatica is that data was stored in an X amount in order to break out the different components of documents in a way that was more easily consumable for computer programs to display later.

Q. And what's the parser?

A. Well, the parser -- in general, a parser would be something that takes a particular stream of data and then extracts the portions of it that have a particular meaning based on their -- their placement in the data stream, for instance.

Q. And the mapper?

A. As I sit here right now, I'm not certain what

the mapper portion does, but given the -- my understanding of the other components, my best guess would be that it -- it sort of relates the data that is parsed out to a more particular format that the rest of the ICCE platform and Informatica software would use to -- to store that data.

Q. And what's your understanding of the serializer?

A. Well, my best guess, as I sit here right now, is that the serializer would be related to actually outputting some of the data that the mapper had transformed into a particular format. That's my best guess given the name of serializer and the way that it appears as the last sub-portion of data transformation.

Q. Have you seen any documents that talk about the mapper or the serializer?

A. I probably have, but I don't recall any right now. This is one such document, for instance.

Q. So you have no recollection as to what these two components do?

A. Well, I gave you my best understanding, as I sit here right now, but I didn't focus on the specifics of how these components worked.

Q. You call those your best understanding or best

guess. Was it a guess, or is it based on information that you reviewed in this case?

A. Well, it's based on their review in this case. As I sit here right now, without other documents in front of me, that's what I can recall.

Q. But you don't remember what documents you learned that from?

A. I don't remember which specific documents.

Q. And did you get that information from somebody at RELX?

A. When you say this -- I'm sorry. Which information are you talking about?

Q. About the mapper and serializer.

A. I'm trying to remember conversations I had with Mr. Groff. I don't think -- I don't think we talked about these two components in detail. They might refer to other things that Mr. Groff and I spoke about. They could have been called different things in those conversations. But I think for mapper and serializer, those were -- my recollection right now is that those were words that were more used in certain documents and not that I discussed with Mr. Groff.

Q. And what's managed file transfer?

A. As I sit here right now, my best guess is that

this is related to actually transferring the files to the Lexis.com and Lexis Advance portions, at least as implemented in the ICCE platform, but I don't -- I don't have a specific recollection of what the managed file transfer process would be in this circumstance.

Q. Could we move to the page called "Our Solution," the second one called "Our Solution." Just a few -- two pages beyond. Not that one. The next one. There you go.

A. Two pages after the last one.

Q. Right. If you look at the "Business Requirements" highlight, do you see that?

A. I do see that box.

Q. Is that your understanding of the business requirements behind the Informatica and ICCE platform?

A. Let me read it. Well, it doesn't seem inconsistent with my understanding of sort of the goals of the ICCE platform and Informatica as a part of that platform for certain time periods.

Q. Do you have an understanding as to whether the Informatica software met those goals?

A. Well, maybe we should go through them one by one. Let's see.

So the first goal is "huge amount of sources

of disparate formats." That seems like a business requirement that reflects more of the sources of data and not the processing, so I don't think Informatica software would have a bearing on that.

Q. Well, wouldn't it have a bearing in the sense that it could actually process a huge amount of sources with disparate formats and do a conversion to a particular format?

A. Right. So for this particular one, I would say that the -- certainly the ICCE platform using the Informatica software was able to process documents for certain time periods that -- that did have different formats, so sure, for this -- this specific goal as stated, there were disparate formats that were processed through the system.

Q. You keep saying "certain time periods." What time periods are you referring to?

A. Well, I'm referring to the time period where Informatica software was a part of the ICCE platform --

Q. Okay.

A. -- and the ICCE platform was processing documents.

Q. Okay. And what about was -- was the Informatica software able to handle huge data volume

individually or that's put together?

MR. SCOTT: Objection.

A. So with this bullet and the last -- and the previous one -- well, just first point out that "huge" is a bit vague. I suppose "huge data volume individually" might mean for individual files, but I'm not really sure as written what that is supposed to mean here. And "or putting together," I suppose that could be for the files in general. But I'm also not sure what is specifically meant by this.

Q. Okay. What about "fast growth on source types and volumes"? Does the Informatica software handle that?

A. Well, the "fast" is also fairly imprecise here. For "source types," I suppose that means the origin of the files, perhaps. I'm not sure how that's distinct from volumes. "Volumes" could mean like a storage volume, but I suppose it could also mean amount of data.

Q. So you don't know?

A. Yeah, I think it would depend on -- on what these words actually are intended to mean. There were a lot of -- well, there were a number of files processed through the ICCE platform using Informatica software. I don't know if the -- for instance, the fast growth here

reflects that or it kind of matches the definition intended by whoever wrote this slide.

Q. Well, you understand that they were LexisNexis employees that gave this presentation, right?

A. Well, their names are on the front of the slide, but I don't know whether they gave a presentation. I don't know whether they authored the slide. The slide has Informatica at the bottom of it.

Q. What about the business requirements of "intricate business process requirements and online on-time goals"? Do you know if the Informatica software met that requirement?

A. I don't know what the requirements are. In this specific goal, it says, "intricate business process requirements and online on-time goals," but it does not state what those intricate business process requirements are or what those online on-time goals are.

Q. Okay. What about the next one, "complex transformation rules and target content models"? Was the Informatica platform able to meet that requirement?

A. So I guess for this particular bullet, "complex" could modify "transformation rules" as well as "target content models." I don't know to what degree complex is supposed to -- well, how do you measure

complexity I guess is my question.

Q. Uh-huh.

A. It's just that these goals as stated are not very precise, so it's hard to say whether or not they were met.

Q. So then the last one, "large user and operational" -- "operation community," do you know if the Informatica software was able to meet that requirement?

A. Well, I don't know what "large" is intended to mean in this context. And I suppose "large" is modifying only "user," but it could be modifying "operation community" in this phrase as well.

Q. Now, given this was a presentation given by two of the people that were heavily involved with the ICCE platform in 2013, do you have your own understanding of what the business requirements were in that time frame for the ICCE platform using Informatica software?

MR. SCOTT: Objection.

A. Well, I don't know if -- I don't know who gave this presentation or whether it was given. Would you repeat that question?

MR. DOYLE: Can you repeat it, please.

(Question read)

A. So in general, my understanding was the -- and

this is mostly from talking with Mr. Groff -- was that the ICCE platform was intended to be a way for RELX to organize its document processing in -- well, in a more organized way. And their goal was to, over time, process more documents through the platform. And they slowly increased the number of -- or percentage of documents that were going through the platform over time.

In terms of the specific goals with respect to maybe how -- well, I don't know to what extent they had milestones for, for instance, we want to process a certain percentage of documents through ICCE versus not, or how quickly to process them through ICCE at certain periods of time. But overall, my understanding was that, in summary, they wanted to use the ICCE platform to process as many documents as they could, but they were going to do it kind of slowly and to make sure things were working as they were going.

Q. And that information came from Dwight Groff?

A. Yes. That's mostly from my conversations with Mr. Groff.

Q. Anywhere else? Can you recall any other documents or conversations that allowed you to make that statement?

A. Well, there was one document in my report that

related to the percentage of documents that were processed through the ICCE platform over time. And in that document, the percentage increased over time to about 15 percent, I think sometime in 2016, if I'm not mistaken, but I'd have to go back and look.

Q. But you didn't know about these business requirements in this exhibit marked 2 that we just went through?

A. Well, I think I've seen this document before, but like I said, these requirements listed here are -- are fairly vague.

Q. Okay. Let's go to the next page. See the third bullet point?

A. Yes.

Q. Can you read it, please.

A. The third bullet point reads: "At a glance, dashboard for monitoring progress of content and adherence to service level agreements."

Q. What is your understanding of the service level agreements?

A. I'm not exactly sure, as I sit here right now, what they refer to, but my guess would be that they refer to a level of service that RELX would provide to its customers.

Q. Measured by what?
A. I'm not sure.
Q. Okay. If we could go all the way to the "Summary" section.
A. About how many pages is that?
Q. It's about four or five I think. It's after the "Our Solution."
A. With "Summary" at the top?
Q. Yep.
A. Okay.
Q. Do you see the "Business Benefits"?
A. Yeah. The first bullet reads "Business Benefit."
Q. And do you understand those to be benefits of the Informatica software on the ICCE platform?
A. Well, these seem to me to be goals. We just went through some previous goals on previous slides, and this, I think, is a summary of them.
Q. Does it say "goals"?
A. Well, it says "requirements" earlier.
Q. And then on the summary it says "Business Benefit," right?
A. Right. It's unclear to me if that is an aspirational benefit or not.

Q. Do you see the first one, "improve content accuracy, completeness, and timeliness"?

A. That's what that bullet says.

Q. Do you know if the Informatica software in the use by RELX improved the content, accuracy, completeness, and timeliness?

A. I'm not sure if it did or did not. That would have to be considered in the context of it operating within the ICCE platform as a whole as well.

Q. So you don't know?

A. I'm not sure, as I sit here right now.

Q. What about the next benefit? Do you know if the Informatica software on the ICCE platform reduced development and operational support costs?

A. I'm not sure what this would be compared to. If you're reducing something, it would be compared to something else. And the same for operational support costs.

Q. So you don't know?

A. Well, I'm not sure what -- what this bullet is even trying to say.

Q. Okay. Do you know what operational support is?

A. My understanding is that operational support

is related to when you have software that is installed and you would like support from perhaps the vendor from whom you bought it to help with its maintenance or administration.

Q. And so what's your understanding of reduced operational support costs?

A. Right. In general, that would just mean that you're spending less money or time or something else that you would consider to be a cost with one implementation versus another, say.

Q. Okay. But you don't know whether or not the Informatica software on the ICCE platform reduced development and operational support costs?

A. Well, I'm -- I'm just not sure what this is compared to. So, for instance, are you comparing it to how the ICCE platform was -- or I'm sorry -- how RELX was processing documents before, or are you comparing it to what RELX might have done an the ICCE platform if they still created the ICCE platform but didn't use Informatica, or are you comparing it to another third-party solution like Informatica but not Informatica itself. So that's -- that's why I'm unclear as to --

Q. Wouldn't this be important to know as

part of your analysis?

A. So my analysis was primarily with respect to the degree to which RELX benefited from having the ICCE platform with Informatica software as part of it, have access to, for instance, 104 cores versus something like 72 or 56.

Q. So benefits are irrelevant to your analysis?

A. That's not what I said.

Q. How do you take into account benefits in your analysis?

A. So, again, my primary -- one of my primary opinions is the degree to which RELX might have benefited from having the ICCE platform with access to Informatica software have access to more cores than they may have been licensed for at different times, but that opinion -- or that -- that idea is distinct from the degree to which RELX benefited at all from having access to the Informatica software.

Q. Did you ever study how RELX benefited at all from having access to the Informatica software?

A. So in general, that wasn't relevant to my opinions because my opinion was focused on the difference in the two circumstances where RELX would have had the ICCE platform have access to Informatica software with

different numbers of cores.

Q. So you didn't -- you didn't study that?

A. Does -- does my previous answer not answer that?

Q. Did you study the benefits of the Informatica software to the RELX platform?

MR. SCOTT: Objection. Asked and answered.

A. So what I studied is the -- is whether there was a difference between the benefit accrued to RELX in the circumstance where the Informatica software was operating as part of the ICCE platform with different numbers of cores, but I didn't look into the specific question of whether RELX benefited in general from the Informatica software.

Q. And in your opinion, you didn't need to know that in order to do your analysis of the benefits between 104 CPU cores and 56 CPU cores?

A. So because I was looking at the difference between what actually happened during two time periods where there were different numbers of cores available for the ICCE platform that included some of the Informatica software, I focused on the data as the record of what happened, and not, for instance, potentially aspirational benefits that were mentioned in a

presentation at some point previous.

Q. You keep saying "aspirational benefits." How do you know it's aspirational?

A. I don't know if they're aspirational or not.

Q. Then why did you just say "aspirational benefits"?

A. Because they may have been aspirational.

Q. They may not have been, right?

A. The benefits listed on this page may or may not have been aspirational.

Q. May or may not, right?

A. The benefits on this page may or may not have been aspirational.

Q. Okay. Good. Go to the next page. Do you see this?

A. I see a page in front of me with the title of "LexisNexis B2B Solution Milestones, May 2013."

Q. Do you know whether this shows that RELX planned growth on the ICCE platform in 2013?

MR. SCOTT: Objection. Calls for speculation.

A. Okay. So I've reviewed the page. Would you repeat the question one more time?

MR. DOYLE: Could you read it back to him, please.

Q. My question is -- well, I'll go ahead and ask it again.

Does this page depict to you that RELX planned growth in 2013 on the ICCE platform?

MR. SCOTT: Objection. Calls for speculation.

A. What do you mean by "growth"?

Q. Do you know what "growth" means?

A. In general, I do --

Q. Okay.

A. -- but I'm asking how you're using it in this question so that I can better answer the question.

Q. How do you use it?

A. Well, I would say growth refers to when something increases in size or some other characteristic.

Q. Okay. Looking at this page, can you determine whether or not anything was increasing in size in 2013 for RELX?

MR. SCOTT: Objection. Calls for speculation.

A. Looking at this page, I am not sure.

Q. Throughout the process from 2010 up until November 2018, did you ever study the -- the growth on the platform and the requirements on the platform in terms of, you know, uploading new content, how fast to upload that new content, what software or machines needed

to be added to manage that content?

MR. SCOTT: Objection. Compound. Confusing.

A. Would you clarify that time range, again? Did you mean 2018 or 2017 at the outset?

Q. 2018. You're right. Thank you.

A. Sorry. Did you mean 2017 instead? I'm confused.

Q. It's 2010 up until November 2017.

A. Okay. Could you repeat the whole question again?

Q. Of course you're going to ask that.

MR. SCOTT: Would you do it, please.

(Question read)

A. So in terms of growth, I did see an increase in the number of documents that were processed on a daily basis through the platform. That was one aspect I agree with that I saw; that was between 2012 and 2017.

Q. Anything else?

A. Let's see. You mentioned -- you mentioned the number of servers or changes in the number of servers that were available for the ICCE platform to access.

Q. All right. Let's move on. Could you go to the -- there's a slide called "Benefits of the LexisNexis Platform Approach."

A. Okay. I'm on the slide that says "Benefits of LexisNexis Platform Approach."

Q. Okay. Do you see down at the bottom where it says "Performance and Scaling"?

A. There is a bullet that says that, yes.

Q. And do you see where it says "minimal always-on workflows"?

A. Yes, where "always-on" is in quotation marks.

Q. Yeah. Do you understand that to be a benefit of the ICCE platform using Informatica?

MR. SCOTT: Objection.

A. Let me review the slide. So this appears to be similar to what we looked at earlier in this same deck. I'm not sure if this is aspirational or not, but this does appear to be about the ICCE platform with Informatica software incorporated given that there's "Informatica" at the bottom of each of these pages.

Q. What does it mean to be always-on?

A. Well, it's in quotes here, so I'm not sure what the author of this document intended it to mean.

Q. Okay. Did you know that a benefit of the ICCE platform with Informatica was to dynamically create unlimited processing workflows?

MR. SCOTT: Objection.

A. Well, the next bullet here does say "dynamically create unlimited processing workflows," but like the other elements of this presentation, I'm not sure if that is aspirational or not. And it also seems a little vague. I'm not sure what "dynamically" means. It could have a number of different meanings.

Q. So you've pointed out a lot in this presentation that's vague to you; is that right?

A. Yes, I'd say there's a number of words here that are not very precise.

Q. Did you ever request to speak to Brian Wisvari or Xin Li who are both the technical -- lead technical engineers for the ICCE platform?

A. I don't recall ever requesting to speak to them directly because I had access to Mr. Groff.

Q. And you think Mr. Groff had all the information associated with the ICCE platform that you needed to know?

A. Well, he seemed -- he seemed able to answer my questions competently or, if not, find a document that did answer them.

Q. Don't you think it's important to verify what Mr. Groff told you?

A. Well, sure. There were a number of things

Mr. Groff told me that were evident in the documents as well.

Q. But you didn't think it was -- could be valuable to speak to Mr. Wisvari or Mr. Li who actually were designing the system --

MR. SCOTT: Objection.

Q. -- and the technical leads on the system?

MR. SCOTT: Objection.

A. Well, Mr. Groff was made available to me as someone who was knowledgeable about the system, and the things he told me were reflected in the documents, so I didn't find a need to talk to anyone about the topics I asked about.

Q. Did -- did that include benefits?

A. Well, like I said earlier, the primary opinion in my report was with respect to the difference in potential benefit of the ICCE platform incorporating Informatica software with different numbers of cores available. I wasn't looking at overall benefit to RELX that it potentially got from installing the Informatica software in general.

Q. Did Mr. Groff tell you how the Informatica software works?

A. He did provide me with an overview of how the

Informatica software integrates with the ICCE platform.

Q. Did he give you information on how the Informatica software works?

MR. SCOTT: Objection. Asked and answered.

A. I do think I just answered that. Is there --

Q. No. You asked -- you answered, sir, "He did provide me with an overview of how the Informatica software integrates with the ICCE platform." I keep asking you did he tell you how the Informatica software works.

MR. SCOTT: Same objection.

Q. Different question.

A. Okay. So in my conversations with Mr. Groff, he explained to me how the Informatica software works within the ICCE platform. My understanding is that the Informatica software did not operate on its own and was instead incorporated into the ICCE platform, which is why I didn't specify that.

Q. How did the Informatica software work?

MR. SCOTT: Objection.

A. Well, we've discussed parts of it already. But in general -- let's see -- the ICCE platform provides a workflow of framework for different components of -- of the ICCE platform. It also provides the data

transformation implementation that is used for some of those workflows, and then there are other implementations that RELX made to operate within that same workflow of framework.

Q. Okay. That's the ICCE platform. What is the Informatica software? How does the Informatica software work?

MR. SCOTT: Objection. Asked and answered.

A. Okay. I'll try to answer the same question again. The Informatica software operates within the ICCE platform. In particular, it provides a workflow of framework for different workflows to operate in the context of processing documents. And then, in particular, there's the data transformation implementation which implements certain of those workflows, and certain other workflows are implemented by custom RELX software.

Q. What's the design of the software?

MR. SCOTT: Objection.

A. Which software?

Q. Informatica software.

A. Which aspect of design are you asking about? I don't --

Q. Any aspect of design.

A. Okay. Well, I don't have access to the Informatica source code, so I can't offer any opinions about how it's designed from a software engineering perspective. In terms of how it operates, my understanding is that it provides a framework for workflows that are utilized by the ICCE platform, at least when the Informatica software was a part of the ICCE platform. And then, for instance, the data transformation component of Informatica implements some of those workflows, and then certain other workflows are implemented by custom RELX software.

Q. Did you read any documents or ask for any documents that would tell you how -- tell you about the design of the Informatica software or how it works?

A. There was one such document that particular I mention in my report, this is the one that was produced through Mr. Groff, and it illustrates different workflows that are -- that are in the system and which workflows had, for instance, the data transformation implementation and which had a Java or Perl implementation from RELX.

Q. Did you read any documents that showed the functional design and components of the Informatica software?

MR. SCOTT: Objection.

A. Well, the document I just mentioned I would say was a functional description of various components, so yes, that was one such document.

Q. And what document's that?

A. It's the one that's in my first report. There's four different -- I forget if I labeled them as figures or not. Well, I guess I can just look at it.

Q. Can you identify it, please.

A. Paragraph 33 of my expert report submitted in May of this year, there was a discussion of a document produced as Expert Discovery 18. That's the Bates number. And then in the subsequent pages of my report, on pages 12 and 13, there are figures that illustrate different pages from that document.

Q. I'm sorry. What page are you on?

A. Page 12 of my report submitted in May.

Q. Okay.

A. 11, the previous page, has a bit of a description; and then pages 12 and 13 have excerpts from the document that I discuss.

Q. And what are those showing?

A. Are you referring to the figures here?

Q. Yeah.

A. Okay. So Figures 1 and 2 relate to ICCE

Platform Version 1. And Figures 3 and 4 relate to ICCE Platform Version 2.0. And Figures 1 and 3 relate to the various workflows that were a part of the ICCE platform. Those were workflows that the framework for which was provided as part of the Informatica software.

And then in Figures 2 and 4, there's a depiction of which of those workflows in the preceding figure were implemented for different types of documents and whether they were implemented by the data transformation implementation as indicated by DT stage, and then a box filled in with DT; or whether they were implemented with a Java or Perl implementation as indicated by a J or a P in those boxes. And those Java and Perl implementations were RELX custom software.

Q. And were all these stages identified in Figure 1, 2, 3, 4, were they all handled by Informatica software?

MR. SCOTT: Objection. Asked and answer.

A. Well, the implementations are in Figures 2 and 3 -- I'm sorry -- 2 and 4, so we could go through them one by one, I suppose.

So for Informatica -- I'm sorry. For the ICCE platform Version 1, Figure 2 --

Q. I'm just -- I'm just asking if all of them

were done by, and performed by, the Informatica software.

A. No.

Q. Is there any more information that you had with respect to the design of the Informatica software other than these figures?

A. Well, this is discussed in my report in paragraphs 28 through 33. We talked about 33, but the preceding paragraphs give some additional color as well.

Q. Okay. When I look at -- let's start with page A.

A. Okay.

Q. It says under A, "The Informatica software is only one component of a complex system that provides functionality of RELX's Lexis Advance and Lexis.com products."

Do you see that?

A. I do see that, yes.

Q. Okay. So on page -- on this page, what does Lexis -- I mean what does the Informatica software do as indicated on this page?

MR. SCOTT: Objection.

A. Well, to summarize, paragraph 29 provides a pretty concise overview of where Informatica software is used in the overall process of getting documents from a

document provider to the Lexis Advance and Lexis.com customers. And as you can see, there are five different steps there. This goes onto page 9. And it is Step 2 where RELX converts the data into a generic data format that takes advantage of the ICCE platform and, for certain time periods at least, the Informatica software that was part of that platform.

Q. Does that tell you how the Informatica software works?

A. Well, I mean, it provides context for the way in which the Informatica software works.

Q. Well, it tells you what it does, right?

A. In my mind, that's -- that's relevant to answering how it works, what it does is its input and output.

Q. So that -- in your opinion, that tells you how the Informatica software works?

A. Can you clarify what you mean by "that" in your question, please?

Q. What you just identified here, under 29, these steps.

A. Well, I'd say it's relevant to how the Informatica software works. I'm not going to sit here and tell you that, you know, a few bullet points are the

entirety of detail that one would potentially write down in terms of how the Informatica software works.

Q. Is that your entire understanding of how the Informatica software works in these bullet points?

A. No.

Q. Well, what's the rest of it? How does it work?

MR. SCOTT: Objection.

A. Well, we can talk about the other paragraphs here that are in my report. I also probably have an understanding that is not explicitly written in this report. I tried to write down the things that were most relevant in this report to providing context for a reader to understand the other portions of it.

I can go through these paragraphs, if you like.

Q. No. I'm just asking you again how does the Informatica software work?

MR. SCOTT: Objection.

A. So I think I've already provided a summary of how the Informatica software works. Would you like more detail on that, or -- I'm just not sure what your question is at this point.

Q. Other than the summary you provided, can you

tell me anything more about how the Informatica software actually works?

A. Sure. I can go into more detail on Figures 2 and 4, for instance. We talked about a number of workflows that are -- that are present. And just referring to these figures, that's Figures 1 through 4 on pages 12 and 13, there are a number of different stages listed in Figures 1 and 3. And the sum subset of these stages are used to process each of the -- the files that go through the ICCE platform. And then some of those stages are -- or workflows are implemented with, for instance, data transformation software from Informatica.

Q. And how is that identified?

A. In these figures? Well, there are a number of stages listed in Figures 2 and 4, and those stages that are implemented with data transformation software have a label of DT --

Q. Okay.

A. -- which in the key refers to the data transformation stage. Those workflows are also described in Figures 1 and 3 in terms of what they actually do in each of those stages.

Q. Okay. Outside of these figures and what you've already testified, do you have any other

understanding as to how the Informatica software works?

MR. SCOTT: Objection.

A. So I have the other paragraphs here in my report in paragraphs 28 through 33 that describe how the Informatica software works. As I sit here right now, I think this is a pretty good summary of it.

Q. Where?

A. Spanning paragraphs 28 through 33.

Q. Do you have any other understanding as to how the Informatica software works?

MR. SCOTT: Objection.

A. I think I probably do given my conversations with Mr. Groff, but as I sit here right now, that's all I can think of at the moment.

Q. Okay. So one of the things you said that you were tasked with was determining whether or not there was any benefit from adding cores up to 104 CPU cores; is that correct?

MR. SCOTT: Objection. Mischaracterization of prior testimony.

A. I don't know if I'd phrase it that way. It is more the extent to which there was a potential benefit for having more cores accessible by the ICCE platform that included Informatica software versus having fewer

cores. I don't think over those time periods that cores were added, for instance.

Q. And in trial, are you going to limit your testimony to that?

A. Well, I have other opinions in my report, so certainly, I'll reserve the right to express those opinions at trial.

Q. And can you summarize what those opinions are?

A. Well, the best summary is the report and the table of contents. I will note that this is only my first of two reports and so there are other opinions in that second report.

Q. Sure.

A. I don't know if it's productive for me to read through the table of contents or to read through the summary of my opinions which are listed at paragraph 9. But that is a pretty good summary --

Q. Okay.

A. -- of the opinions expressed in this report.

Q. What I'd like to do is just go through and talk about what actually happened here between the parties of RELX and Informatica and get your understanding of those.

So you're aware that on May 28, 2010, the

parties signed a master service license agreement?

A. I believe there was a such agreement signed in May 2010. I don't recall the specific date.

Q. Okay. And are you aware that the number of CPU cores licensed was limited to 16?

A. That matches my recollection as I sit here right now, though I think that was also before processing had actually begun on the production ICCE platform.

Q. And do you know how it was determined by RELX why they needed to license 16 cores at that time?

MR. SCOTT: Objection.

A. I don't know why RELX decided to license 16 cores.

Q. Okay. In December 20, 2011, were you aware that RELX added four cores to the license?

A. That matches my recollection. I'm not sure of the date of that agreement.

Q. Why did they add four cores?

MR. SCOTT: Objection.

A. I'm not sure why RELX decided to add those cores.

Q. Okay. Are you aware there was a 2012 amendment dated June 29, 2012, and during that amendment, 52 new cores were licensed? Are you aware of that?

A. My recollection is there was an agreement like that. I -- I'm not sure whether the parties dispute the number of cores that were available at that time, but I do recall that there was an agreement that increased the number of cores.

Q. And after that agreement, the aggregate cores license was 72? Do you recall that?

A. I believe that number might be in dispute, but I understand that's -- that's one way that document could be read.

Q. Okay. Why did RELX license 52 new cores in the 2012 amendment?

MR. SCOTT: Objection.

A. I don't know why --

MR. SCOTT: Speculation.

A. I don't know --

Q. Why would they do that?

A. I don't know why RELX decided to license more cores at that time.

Q. Did you ask anybody?

A. I don't recall that I did.

Q. Did you look for that information in any documents?

A. Well, the focus of my analysis was on what

actually transpired with respect to how many cores were available at different time periods, so --

Q. Did you look for that information in any documents?

A. I did not look for why RELX made certain decisions that they made.

Q. Okay. Are you aware that as of March 31, 2013, RELX was actually using 104 CPU cores?

MR. SCOTT: Objection.

A. My understanding is that around that time period -- and I don't recall the specific date, but around that time period, there were 104 cores that were in the ICCE platform on servers that were accessible by the ICCE platform including Informatica software. I didn't know -- there is no data that indicates that individual cores were used at certain time periods or not.

Q. Do you know why RELX deployed Informatica software on 104 CPU cores?

MR. SCOTT: Objection. Assumes facts not in evidence.

A. I think there's a dispute about how the Informatica software was actually deployed at certain times.

Q. I'm not asking that.
A. Would you repeat the question, then?
Q. Yeah. I'm asking you do you know why RELX used 104 CPU cores --
MR. SCOTT: Same objection.
Q. -- with the Informatica software installed on it?
MR. SCOTT: Same objection.
A. So I don't know why RELX did certain things. But, again, I don't think there is any data to show that all 104 cores were used at certain times or not. My understanding is that around that time period, there were 104 cores available for the ICCE platform that incorporated Informatica software, but I don't know exactly which cores were used by the software at what times.
Q. But why did they -- why did they add cores to get up to 104 CPU cores?
MR. SCOTT: Objection. Calls for speculation. Asked and answered.
A. I don't know why RELX made the decisions that they made.
Q. Did you ask anybody about that?
A. Since I was focused on the data related to

utilization of the -- of the servers at those times, I didn't focus on -- on why RELX made certain decisions.

Q. Do you think they received a benefit from going from 32 CPU cores up to 104 CPU cores? Might that be the reason why they spent a lot of money and added all those cores?

MR. SCOTT: Objection. Calls for speculation. Asked and answered.

A. So I didn't look at that specific question of that increase of cores that were available to the ICCE platform.

Q. You didn't think that was relevant, right?

A. Well, the question I was focused on is whether -- for a later time period, whether -- or I'm sorry -- for the time period during which there were a number of cores that were available to the ICCE platform that was greater than the number of cores that were licensed, whether or not there was a benefit accrued there.

And could I also request a break some time soon?

Q. Sure. I've just got a few more questions on this sheet.

So you're aware, I think we talked about this

earlier, that in February 11, 2015, notwithstanding the fact that RELX had Informatica software on 104 CPU cores, they went and negotiated to retire 16 CPU core licenses. Do you recall that?

MR. SCOTT: Objection.

A. Well, I wouldn't -- I think you said the Informatica software was installed on CPU cores?

Q. I didn't say that.

A. Well, could you --

MR. DOYLE: Do you want to read it back to him.

(Question read)

A. Okay. So in terms of Informatica software on cores, I'd say it was installed on the servers. But disregarding that -- that point, I don't know why RELX decided to decrease the number of licenses that they -- that they had with Informatica during that time.

Q. Do you know whether they wanted to get some money back?

MR. SCOTT: Objection. Calls for speculation. Asked and answered.

A. I don't know why RELX would have made that decision.

Q. Do you think that they told Informatica they

didn't need 72 CPU cores to process the information and all they needed was 56?

MR. SCOTT: Same objection.

A. I don't know why RELX decided to make that decision one way or another.

Q. Well, isn't that similar to what your opinions are in this case: that RELX only needed 56 or fewer cores to process the Informatica software?

MR. SCOTT: Objection.

A. Well, the question of why they made a decision I think is distinct from whether or not they actually needed a certain number of cores.

Q. Well, there could be relevance, though, in terms of why they made certain actions by adding or decreasing cores. That would have impact on your analysis, sir. Wouldn't you agree?

MR. SCOTT: Objection.

A. Well, what RELX decided -- I mean, they -- they may have had imperfect information or imperfect reasoning, so for the purposes of my opinion, I was focused on the actual data that reflected what actually transpired during those time periods.

Q. So you say they may have had imperfect reasoning. And what makes you to believe that they

had imperfect reasoning?

MR. SCOTT: Objection. Mischaracterization of testimony.

A. I think I said they might have had misinformation or have reasoned in various ways, but I don't know one way or another why they made certain decisions.

Q. Well, is there anything that makes you believe that they had imperfect reasoning when they kept 104 CPU cores installed with Informatica software but yet negotiated to decrease the license amount from 72 to 56?

MR. SCOTT: Objection. Compound. Misleading.

A. Like I said, I don't know why they would have made certain choices. If that doesn't answer the question, could you repeat it? I'm sorry. It's a bit long.

Q. Did you ask Dwight Groff, who it seems like you got a lot of information in this case -- did you ask him why they negotiated to remove 16 CPU core licenses?

A. Since I focused on the data in this case, I didn't ask about why RELX made certain decisions.

Q. Did you ask Mr. Groff or anybody else why, from approximately March 31, 2013, all the way up until mid -- to May 2016, why they continued to operate with

Informatica software on 104 CPU cores?

MR. SCOTT: Objection.

A. Well, assuming those dates are correct, I don't recall the specific dates, but, again, I was focused on the data and what the data actually reflected in terms of what transpired, so I didn't ask questions with respect to why RELX made certain decisions.

Q. Do you know whether they did that in order to achieve better performance --

MR. SCOTT: Objection. Asked and --

Q. -- on the platform than merely using 72 or 56 CPU cores?

MR. SCOTT: Objection. Asked and answered.

A. Like I said, I don't know why they made that decision one way or another.

Q. Could that have been one of the reasons?

MR. SCOTT: Objection. Calls for speculation.

A. Would you repeat that? Sorry.

MR. DOYLE: Can you repeat the --

Q. Well, here, I'll just go -- I was asking you whether or not the reason why they continued to use 104 CPU cores when they were only licensed for either 72 or 56, that the reason they continued to use those is because of certain performance factors.

MR. SCOTT: Objection. Calls for speculation.

Q. Could that be the case?

MR. SCOTT: Same objection.

A. I don't know why they decided to do the things that they did. There could be reasons related to perceptions about the performance, I suppose, but like I said, I don't know why they made choices one way or another.

Q. Why didn't you ask them to see -- to see? Wouldn't that make your job easier?

MR. SCOTT: Objection. Asked and answered.

A. Well, like I said, I was focusing on the data because that's the actual record of what transpired. Perceptions notwithstanding, that's --

Q. Well -- well, yeah, you focused on the data, the data -- one of the data being CPU utilization, right?

A. That is some of the data that I looked at, yes.

Q. Do you know if RELX employees actually had access to CPU utilization charts going all the way back to the beginning of this in May 2010?

A. My understanding, and I addressed this in my report, is that the first time that RELX pulled records for CPU -- CPU utilization, they only had records going

back two years because that data was overwritten after two years. And so initially, they provided two years prior to the first time that they looked into it, and after I requested the data for after that time period, they also provided that data as well.

Q. And what was the two years that they had access to?

A. I'll have to look back in my report. I think it was -- it was approximately -- let's see. So the data afterwards was approximately November 2017 to, I think, something like June 2016, and so I believe it was two years prior to that, so let's say approximately mid-2014 to mid-2016.

Q. So it's your understanding that prior to mid-2014, they did not have access to CPU utilization data?

A. Let me clarify one thing and then answer that question. One moment. Okay. So in terms of what data was available with respect to utilization data, my understanding is that the earliest time period at which data exists begins December 9, 2014.

And to answer your question specifically with respect to whether data prior to that existed, it's my understanding that that data did exist at one point in

time, but at the time at which the data was requested because of this litigation, at the time that data was requested in late 2016, the data prior to December 9, 2014, had already been overwritten and was no longer available.

Q. So there was data prior to December 9, 2014, on CPU utilization; is that right?

A. My understanding is that at some point that data existed, but at the time this litigation began, the data no longer existed.

Q. What happened to it?

A. My understanding is that RELX's policy was to retain CPU utilization data for two years, and then after two years, it would roll over, and so they would always have the most recent two years but not data prior to that.

Q. And when was that data destroyed?

MR. SCOTT: Objection.

A. Well, I'd say it was overwritten. And the way I understand it would work is -- so on every day, there is data recorded for -- for that day, and then the day that is two years and one day prior would be overwritten by the previous day's data.

Q. Okay. So the data up to December 9, 2014,

would have been overwritten when?

A. It would have -- my understanding is it would have been overwritten, I guess, just on or around December 9, 2016. The idea being that they always have the most recent two years of utilization data, but they don't store it beyond two years.

Q. So prior to December 9, 2014, they did have in existence CPU utilization data?

MR. SCOTT: Objection. Asked and answered.

A. I'll just remind that I'd like a break soon.

But with respect to data, so it's my understanding that data prior to December 9, 2014, did exist at one point in time, but as of around December 9, 2016, that data no longer existed.

Q. Okay. And so let's go all the way back to December 2014. Is it your understanding that there was CPU utilization data in existence at that point?

MR. SCOTT: Objection. Asked and answered.

A. It's my understanding that given the way the data was retained at RELX, that at any point in time, there were two years of data available that is the most recent two years of data on a particular date.

So, for instance, supposing it was December 5, 2014. At that point in time -- I'm sorry -- December 5,

2014, is what I meant to say -- there would have been data for the previous two years available, so that would be December -- approximately December 5, 2012, to the current date at the time which is December 5, 2014.

Q. Okay. And it would also be your understanding that prior to December 5 -- or December 2012, there also would have been CPU utilization data that was available to RELX, right?

MR. SCOTT: Objection.

A. I'd say I'm probably a little bit less confident about that because that was further in the past. I don't know exactly when they started retaining CPU utilization data in the way that they did in the 2014 to the 2016 time frame.

Q. Okay. So from this time period from 2010, let's say, up until -- well, up until 2016, do you know whether RELX was looking and using the CPU utilization data?

A. It's my understanding from talking with Mr. Groff that RELX employees would occasionally look at it, but generally, only if there were some problem with the operation of the software to help diagnose the problem.

Q. Do you have any other means other than

Mr. Groff who told you this?

A. As I sit here right now, I can't think of any other -- any other conversations about that topic.

Q. And when did Mr. Groff start employment with the company -- or when did he start being a consultant to the company?

MR. SCOTT: Objection.

Q. What's the first date?

A. I don't recall the exact date where he started working for RELX.

Q. Well, was he at RELX in 2013?

A. Like I said, I don't recall the exact date.

Q. 2014?

A. Like I said, I don't recall the exact date.

Q. Well -- okay. So you said --

MR. SCOTT: Let's take a break.

MR. DOYLE: Yeah, one -- I'm going to finish this. I've got two questions.

Q. So Mr. Groff -- you said that Mr. Groff told you that employees at RELX would look at the data but only if there was a problem. What time period was he referring to?

A. Well, I don't know if he limited it that strictly. My recollection is that he said generally they

don't look at the data but they usually look at it if there's a problem. And then -- sorry. What was the question?

Q. Over what time period?

A. Is your question over what time period they would look at the CPU utilization data?

Q. Yeah.

A. I don't think his answer was limited by time. I think he was just talking about the utilization data in general, so for the time periods during which it existed, I imagine.

Q. So even before his employment or association with RELX?

MR. SCOTT: Objection.

A. I'd be speculating here because I don't recall exactly what he said, but my understanding of the conversation is that he was talking about in general how RELX employees typically looked at the utilization data, and so it wouldn't surprise me, for instance, if they were looking at it for all time periods during which it existed if those time periods preceded Groff's employment there or not.

Q. And do you know whether or not they looked at CPU utilization data when they made these decisions to

add more licenses? For example, when they went from 20 CPU licenses up to 72 CPU licenses, do you know whether or not they considered CPU utilization data?

MR. SCOTT: Objection.

A. I don't know whether or not RELX considered utilization data when coming to that conclusion, or any other data coming to that conclusion.

MR. DOYLE: Okay. Take a break.

THE VIDEOGRAPHER: The time is 3:19 p.m. We are off the record.

(Recess)

THE VIDEOGRAPHER: The time is 3:43 p.m. This is the beginning of Disc 4. We are on the record.

Q. Sir, do you know that when RELX added CPU licenses to go up to 72 CPU cores, do you know whether or not that that -- those additional CPU cores were added based on CPU utilization charts?

A. So I don't know why RELX made that decision. I don't know if it was based on CPU utilization charts or other information.

Q. But it could have been based on CPU utilization charts, right?

MR. SCOTT: Objection. Calls for speculation.

A. It could have been based on any number of

things, I suppose. CPU utilization data could have been one of those things, but I don't know one way or another.

Q. And when actually RELX used the Informatica software on 104 CPU cores, that decision may have been based on CPU utilization charts, right?

MR. SCOTT: Objection.

A. So I don't know whether the Informatica software was used on any specific cores at any particular time. My understanding is that the Informatica software was -- was installed on servers that had available to them 104 cores as part of the ICCE platform processing. The decisions about whether to reduce or increase licensed cores, that's something I don't know why RELX would have made those decisions.

Q. But they may have made those decisions, at least in part, looking at CPU utilization charts, right?

MR. SCOTT: Objection. Calls for speculation. Asked and answered.

A. So I don't know why RELX made certain decisions. It could have been based on a number of factors. One of those factors could have been CPU utilization charts for certain time periods.

Q. Okay.

MR. DOYLE: I'd like to introduce to -- have

the court report mark as Rucinski Exhibit 3 a series of Excel spreadsheet screenshots that were part of Expert Discovery 000008.

(Exhibit No. 3, Expert Discovery 000008 Excel spreadsheet marked for identification)

Q. Mr. Rucinski, are you familiar with these -- this Excel spreadsheet and the information that's depicted on these pages?

A. Having looked through this, this appears to be the spreadsheet produced in this case as Expert Discovery 8; and yes, I have reviewed this document before.

Q. And so is this -- let's look at the top one first, if we could. Is this a chart that's in your expert discovery?

MR. SCOTT: Objection.

A. You're asking whether this chart appears in my report?

Q. Yeah.

A. No, this chart does not appear in my report.

Q. Did you base any information on this chart?

A. On this specific chart, no.

Q. Did you use charts that were similar to this?

A. Well, there were charts in my report that were based on data that was present in Expert Discovery 8,

data present in tabs that are not, for instance, aggregate graph as depicted on this first page.

Q. Okay. Now, looking at the first page -- now, is this based on the same data, if you look at this page, that was used in the charts in your report for Figure 12 and 13?

MR. SCOTT: Objection.

A. So the way I would phrase it is that the chart in Exhibit 3, which I have been handed, the chart on that first page I believe was created, in part, based on other data in Expert Discovery 08, and I used some of the same data in generating Figures 12 and 13 in my report.

I just wanted to add also that I didn't scrutinize the -- the chart at the top on the first page of Exhibit 3 that you've handed me because I wanted to use the underlying data directly instead of relying on charts that were produced as part of Expert Discovery 8.

Q. So what data did you explicitly use?

A. Well, there was data in the tabs for each of these seven servers, so that would be -- some of them are cut off here -- but that would be psc3813, psc3814, psc3815, psc3816, psc33817, psc33841, and psc33842.

Q. Okay. And the date -- the data shown in this particular chart is from the Aggregate Data tab; is that right?

A. I'm not sure about that. Like I said, I didn't look at this chart in detail. I don't see the source in this exhibit, but maybe it's on one of the other pages, if you could point to it.

Q. Well, before we go there, did you use data that's from the Aggregate Data charts or from individual servers?

A. No, I used the data that was specifically from the tabs for the individual servers.

Q. Okay. Can you turn the page, please, sir.

So I'm just trying to figure out how this works. So if you look at this particular Excel spreadsheet, in the top, sort of near the top, is this information -- if you look here, it says 33 -- psc33841. Do you see that?

A. I do.

Q. And does that show that data that specified -- you can see there's actually a bar around it. It looks like somebody selected it in Column I, comes from the tab psc33841 at Column F2?

A. My understanding is that that data in Column I

in this chart would come based on -- what's in this text box would come from Cell F2 in the chart psc33814 in that tab.

Q. Okay. So when it says -- when you go over there and it says "psc33841 cores," and it's highlighted "2.4256," what does that refer to?

A. If I'm reading this printout correctly, that should be the same number that appears in Cell F2 of the Tab psc33841.

Q. Is that an amount of data or what is that?

A. Sorry. Would you repeat that? I didn't hear it.

Q. Well, I just -- I'm trying to figure out what that indicates, the 2.4256?

A. So what quantity is that, you're asking me?

Q. Yes, what is that?

A. Well, I'd have to look back at that cell in psc33841. Is it on the next page?

Q. Well, I've got to see if that's the same day. That's a different day.

So you don't -- you don't know what that 2.4256 represents?

A. Well, I'd want to look back at that specific cell from the Tab psc33841.

Q. Okay. Let's go to the next page. Have you seen this document before?

A. Yes. This appears to be the Tab psc33841 in the spreadsheet Expert Discovery 8, specifically the sum of the rows related to June 15 and 16.

Q. Okay. And let's see. Do you see where it's, I guess, the psc33841 is identified, right?

A. Yes, that tab is selected at the bottom, so I presume that the table here is from that tab.

Q. Okay. And so when you look at that, the highlighted portion, let's just say we're looking at the first two items in 6/15, does that refer to seven and eight o'clock in the morning, or the seventh hour and the eighth hour of the day?

A. I think more accurately it describes the eighth and ninth hour. The reason for that is because the hours start counting at zero, and so you can see --

Q. Okay.

A. -- that penultimate row for June 16, 2017, there's a zero there.

Q. Now, let's look to the right of that. Do you see an E? It says 104.29 for the first entry, and then the second one in that row is 10 -- I'm sorry -- column is 106.72. Do you see that?

A. In Column E, the first two cells are 104.29 and 106.72.

Q. And what do -- what do those indicate?

A. Well, there's no heading for the column here, but my recollection is that this -- the numbers in this column indicate the -- well, let me flip for a moment. My recollection is that those refer to the average CPU utilization across all the CPUs for this specific server psc33841 for the two hours here, but I would -- I'd prefer to see the column heading which isn't here.

Q. Okay. Well, if that was the capacity, does that show that they're over 100 percent at what's been indicated as the eighth and ninth hour, that that server was over 100 percent?

A. So the way I'd phrase it is for those two hours, the eighth and ninth hour of June 15, 2017, for the eighth hour, on average, the CPU cores for the server were at 100 for approximately 104 percent utilized and 106 percent utilized for those two hours respectively.

Q. So does that mean for that server that all 32 cores used were near maximum capacity?

MR. SCOTT: Objection.

A. I guess it depends on what you mean by "near." But I do think it's the case that for these two specific

hours, given that the average utilization of all of the cores for the server exceeded by a small margin 100 percent, I'd say that the cores in that server were at least close to fully utilized for -- for those two hours.

Q. Okay. Could we go to the next page, please. Have you seen this page before? And it seems to relate to the previous page, does it not?

A. It relates in that this is a different tab of Expert Discovery 08. This tab is -- is called "Raw Data." And when we compare it with the previous page, it appears that this has the same data as on the previous page. This time there is a column present that says CPU percentage, and this does appear to be the same data from the previous page of Exhibit 3.

Q. And, again, it shows for the same two hours, you're over 100 percent; is that right -- with respect to psc -- pcs -- I'm sorry -- psc33841?

A. So I'd say for those two hours, this indicates that, on average, the CPU cores of that server were more than 100 percent utilized.

Q. And doesn't that mean so all the 32 cores were used or near maximum capacity or at maximum capacity?

A. Just like with the previous page, given that this is the same data, it does seem that because the

average CPU utilization for -- for the cores in this machine for these two hours is above 100 percent, I would say that probably the cores of that machine, the 32 of them, were at least near full capacity for those two hours, though, I don't know what else was going on on the other machines during this -- during these two hours, given this page in front of me.

Q. So if we go to the next page, I believe the next page is the Aggregate Data tab, is it not?

A. The next page does appear to be an excerpt from the Aggregate Data tab, the Expert Discovery 8.

Q. Okay. If you look at the same two hours that were specified on the two previous as being seven and eight, do you see that?

A. Yes, I do see that.

Q. Okay. And do you see for the psc33841 cores, it shows 3.1072 and 10.1056? Do you see that?

A. I do see that.

Q. So in terms of the total cores that are used here, why are the cores here less than 32?

A. Give me a moment. So for this particular page, there's -- well, let's talk about these one at a time.

So there's a cell that appears to be

highlighted with a value of 3.1072. The calculation for that cell, as depicted on this page, indicates that it was -- that this value seems to be taken from the Tab psc33841 in the Cell F22038. So the next thing I would do to determine the origin of this number is look at that cell in Tab psc33841. In this exhibit, I don't think I have access to that cell as labeled. Two pages prior, there's an excerpt but -- for Column F, but I don't see the row numbers on this page.

Q. Why -- why are the cores depicted as only 3.1072 and 10.1 --

MR. SCOTT: Objection. Asked and --

Q. -- for 33841 here?

A. I'm sorry. Did you say 33841?

Q. Yeah.

A. Where are you looking at?

Q. On the last page. If you go over to I, see where the -- Column I?

A. Did you say Column I?

Q. Yeah.

A. 33841?

Q. Yeah. Do you see ps -- at the top psc --

A. Oh.

Q. -- 33841, right? Do you see that?

A. Sorry. I thought that was the number in the cell, not the --

Q. And that's the same server we've been looking at for hour seven and hour eight on June 15, 2017, right?

A. Yes, those are the same hours, so I don't know why these numbers are different. I would have to look at that particular cell indicated here at Cell F2238 --

Q. Why are the cores -- why are the cores here less than 32?

A. Let me finish. I'll just note again that I didn't rely on the aggregate data in coming up with the charts that are in my reports. And -- sorry. Would you repeat that question again?

Q. Yeah. Why would -- why -- why does it have 3.1 and 10.1 instead of 32?

MR. SCOTT: Objection. Asked and answered.

A. I'm not sure. I would want to look at Cell F22038 on the Tab psc33841 in order to determine that.

Q. Did RELX rely on aggregate data at all?

MR. SCOTT: Objection. Calls for speculation.

A. I don't know whether RELX relied on this particular Expert Discovery 8 that was produced in this litigation. I don't know whether they relied on it for anything.

Q. Who produced these charts?
A. I believe RELX did.
Q. Did they develop these charts for the litigation?
MR. SCOTT: Objection. Calls for speculation.
A. By "charts," do you mean the tables that we're looking at here?
Q. Yeah, the spreadsheets.
A. So my understanding is that these specific spreadsheets were produced for this litigation but that the data existed in their ordinary course of business.
Q. But they weren't specified in spreadsheets like this? The data wasn't?
A. I don't know exactly how it was specified in the ordinary course of business.
Q. Could it have been in these spreadsheets?
MR. SCOTT: Objection.
A. Well, I don't know, but I doubt that it would be in a spreadsheet that was labeled with a Bates number for this litigation.
Q. And what do you mean by the "Bates number"?
A. I'm referring to Expert Discovery 8 --
Q. Okay.
A. -- which is the file name for this document.

Q. Who provided you these charts -- or the spreadsheet?

A. I got them directly from counsel.

Q. And what was the data they used to prepare these spreadsheets?

MR. SCOTT: Objection.

A. It's my understanding that this data came from records that RELX had about the average CPU utilization for servers on an hourly basis as reported by the operating system for those servers.

Q. Do you know who generated these spreadsheets?

A. My recollection is that it may have been some combination of Dwight Groff and Jeff Hoffman, but I don't know the exact way they were generated.

Q. You don't know the exact what? I'm sorry.

A. The exact way that they were generated.

Q. Well, I'm not asking the way right now. I'm just asking who generated the charts.

A. As far as I understand, Dwight Groff and Jeff Hoffman had some involvement in generating them. There may have been other people as well. I'm not sure.

Q. Did you look at the original data?

A. When you say the "original data" --

Q. That populates these charts.

A. My understanding is that that data is present in these charts, and so I looked at that data here in these charts.

Q. Did you look at the data -- the original data prior to it getting into these charts?

A. Well, there was some data that was produced as another document with a different Bates number. That would -- that was the -- the utilization data through 2016, I think December 2016. So I looked at that data before looking at the data in this more complete chart which had data through November 2017.

Q. And did you satisfy yourself that that particular data, the origin data, was properly input into these charts?

A. Well, my understanding is that it was the best data available, and I didn't see any material discrepancies with respect to anything else I had available to me, so this data seemed relatively complete, and I had no reason to doubt it.

Q. No. What I'm asking you is actually the source of data, the source documents for the data that somebody used to take the data and then included in these charts. Did you review the source data in whatever format it was in?

MR. SCOTT: Objection.

A. Well, I reviewed that data as it appeared in these -- in these charts. I didn't review the data in a form other than the two charts that I mentioned that was produced in this litigation.

Q. How did you satisfy yourself that it was correct and accurate?

A. Well, having reviewed it manually, and having noted some of the omissions in the data, I felt confident that there was enough of it here. And given Mr. Groff's statements about how this was the best data that they had, it seemed like data that was -- that was complete enough in order to do my analysis.

Q. So, again, you're relying on Mr. Groff?

MR. SCOTT: Objection.

A. It's my understanding -- well, from speaking to him, that this is the best data available, and, you know, he claimed that this was the best they could do, and this came from the operating systems as the operating systems directly reported the average CPU utilization, and that story made sense to me. I had no reason to doubt it.

Q. Did you do anything to verify the data in this spreadsheet on your own?

A. Well, I looked through it myself, and as I mentioned earlier, I did some -- some analysis in Excel just to make sure -- or at least account for the places that there was no CPU utilization data reported. So I did scrutinize it to make sure it made sense and, you know, make sure that I accounted for any -- any omissions.

Q. So we had discussed how, for this particular server, at the hour of seven and eight, it appeared to be, you know, at over 100 percent capacity, which you had testified that each of the cores would be at least somewhat near capacity. But when we look on this last sheet, why does it show such low number of cores --

MR. SCOTT: Objection. Asked and --

Q. -- being used?

MR. SCOTT: -- answered multiple times.

A. So for the sheet, what I'd want to do is look at the origin of the data here which, for this particular cell that's highlighted, appears to be Cell F22038 on Tab psc33841.

As I sit here right now, I don't know why that number appears in that cell, but the next thing I would do would be to look at the cell that it's indicated that it originated from on this printout.

Q. So you would -- what would you do? Click on 3.1072 and see what that provides you?

A. Well, in this table, I think that's already been done. So that cell is highlighted, and, then, in the text box at the top, it indicates the cell that this data is coming from, and so my next step would be to look at that specific cell that is the Cell F22038 in Tab psc33841.

Q. Do you have an opinion as to the data that we're looking at right now with respect to that server whether it's correct or not?

MR. SCOTT: Objection.

A. Well, the data we're looking at right now, I would say is the data on the Aggregate Data tab. Like I said earlier, I didn't rely on the data in this tab. I relied on the data for the specific servers. So I didn't scrutinize the data on this tab. I'd tried to get to the most direct and original source of the data in my analysis.

Q. Sir, could you please go to paragraph 55 of your report.

A. Okay. I'm at paragraph 55.

Q. Now, if you look at Footnote 28 --

A. Yes.

Q. -- is that the math you used for your sizing estimate?

A. Let me review for a moment. So in Footnotes 28 and 29, I was explaining the math that I used to come to the number of 3.2 gigabytes per hour for this particular set of data that I inputted into the sizing model for RELX as discussed in paragraph 55 of my report.

Q. Okay. So looking at Footnote 28, for example, did you take the total number of files and divide it by the total days and multiply by 150 kilobits per file? Is that the right way to read it?

A. That's true except it's kilobytes and not kilobits.

Q. Okay. And is it true that that math would give you 533,212 files per day?

A. I'm sorry. Did you say "files per day"?

Q. Yeah.

A. Well --

Q. Or would it give you kilobytes per day?

A. So that whole expression in Footnote 28 -- let me look at it for one moment -- that would give you kilobytes per day.

Q. And would that be 533,212? Do you know?

A. I believe it would be, as is in my report,

79,981,929 kilobytes of data processed per day.

Q. If you divide 974,179,899 files by 1,827 days, do you end up with 533,212 files per day?

A. I'm happy to do long division for you.

Q. Sure.

A. Okay.

Q. Do you have a calculator by chance?

A. I don't. I left my phone elsewhere. Can I have a pen as well?

Q. Sorry to make you do that.

A. I don't remember the last time I did long division by hand. Okay.

MR. SCOTT: This is comical. Here's a calculator.

THE WITNESS: Oh, thank you.

MR. DOYLE: I'm not sure I trust the calculator, Counsel, but I'll let you do it this time.

MR. SCOTT: You can double check it, Scott. Since I'm not operating it, you got a much higher accountability that it's going to be right.

MR. DOYLE: Well, maybe that's been the problem the whole time: your calculator's off, though. Not you, the calculator.

MR. SCOTT: The calculator is fine; the

"calculat-or" is bad.

A. Okay. So if I inputted that correctly, 974,179,899 divided by 1,827 gives us approximately 533,213.

Q. Okay. So 533,213 files per day; is that right?

A. I believe that's correct, yes.

Q. Okay. Thank you. And then you multiply that by 150 kilobits -- kilobytes per file; is that right?

A. Yes, that's the calculation in Footnote 28.

Q. And did you always use 150 kilobytes per file in your calculations?

A. Well, I used it in this calculation.

Q. Did you use it in others? Do you know?

A. As I sit here right now, I think it was only in this paragraph that I used that number. Though, I'll note that because that 150 kilobytes per file shows up in the Footnote 28, and then that result is used in Footnote 29, that number also affects the result of Footnote 29.

Q. What does 150 kilobit -- kilobytes per file represent here?

A. So this represents -- in the e-mail from Mr. Groff, which we looked at earlier today, he mentioned that across all of the files that were processed by the

ICCE platform, his best guess for the average size of those files was 150 kilobytes.

Q. So did you use 150 kilobytes per file in your calculations?

A. Well, I did use 150 kilobytes per file in Footnote 28, for instance.

Q. Did you use it in other calculations?

A. Are you asking did I use it or did Mr. Groff use it?

Q. Did you use it? I mean, who did the math here? Is this your math or Groff's math?

A. No. I did this math.

Q. Okay. Did you use it in other calculations?

A. Well, as I sit here right now, I think it's -- I think it's just the math that's in this paragraph, so that includes Footnote 28 and 29.

Q. Does that represent an average file size?

A. It does represent an average file size, yes.

Q. Over what period of time?

A. My recollection is that it's for the entirety of the files that were processed through the ICCE platform.

Q. For the entirety of the files processed through the ICCE platform?

A. That's my recollection, yes.

Q. Over what period of time?

A. Well, my recollection is that the files were processed through ICCE platform starting November 2012 and ending in November 2017.

Q. And so my -- I'm going to ask my question again. Do you know if this 150 kilobytes per file, did you use that in other calculations other than the one you have in 28?

MR. SCOTT: Objection. Asked and answered.

A. As I sit here right now, I don't remember other calculations besides Footnote 28 and 29 that would have relied on that number.

Q. Is this the only number that you ever used for the average file size?

A. There was at least one other instance in my report where there was -- there were other estimates of files. Well, hold on. Let me think about that for a moment. The instance I'm thinking of was related to the average amount of data that was processed through the ICCE platform on an hourly basis, so it's not exactly an average file size, but it would be related to it.

Q. And where was that?

A. One moment. So, for instance, in paragraph 54

of this first report of mine just back a few pages, I discuss information I got from Mr. Groff that I used to populate one version of the document I refer to as Informatica Sizing Model for RELX, and those values are depicted in Figures 8 and 9 on the subsequent page.

Q. Can you point me to where those figures are?

A. They're on page 29.

Q. Okay. Where in the charts are they?

A. So the number I'm thinking of in particular is 0.8, and that is in Figure 8 next to the label number of gigabytes per hour.

Q. Okay. And what does that represent?

A. So the context of this is I had asked Mr. Groff to provide me with his best estimates for the number of gigabytes that were processed by the ICCE platform on an hour-by-hour basis over the time period that the ICCE platform was processing files. And I also asked him to break that down by different time frames within that overall time frame if -- if that estimate would change for different times.

Q. So --

A. So -- well, let me finish. So for the specific number, this was the highest average gigabytes per hour that Mr. Groff specified.

Q. And what time frame did this cover?

A. I would have to look back at Exhibit H. One moment. So referring to Exhibit H in this report of mine that I submitted in May, there are -- the first six pages of Exhibit H are the contents of the e-mail, and then after, as part of the exhibit, are three tabs of a spreadsheet which Mr. Groff returned to me as an attachment to this e-mail. The second of those tabs is labeled March 2013 through March 2015, and this is the spreadsheet where he indicates a number of gigabytes per hour of 0.8.

Q. Okay. And what did you do to verify that those numbers by Groff were actually accurate?

A. Well, I wasn't sure how accurate they were, so in performing this part of the analysis in my report -- let me just get the paragraph numbers. So this is Section 8 of my report. It's spanning paragraphs 49 through -- well, let's say it's spanning Section 8 point -- or 8 and then Subsection A. So that's spanning paragraphs 49 through 55.

So what I did here is I was looking at the document I referred to as Informatica Sizing Model for RELX, which is a document produced by Informatica that appears to estimate the number of CPU cores that

installation might require. And I approached this by gathering data from different sources that could be used to populate this document that would calculate the number of cores, and I tried to find, as best I could, whatever reasonable estimates would provide that the largest estimate out -- the largest calculation out of this document, which I found to be 12 after going through a few different ways of populating the document data. So requesting this information from Mr. Groff was one of those ways.

Q. And what sources did he gather from?

A. I'm not sure what sources he used in order to reply to my e-mail.

Q. Okay. Let's go back to your e-mail, if we can.

A. Okay. I'm at Exhibit H.

Q. Okay. So if you go to the page 2 of Exhibit H, is that your question, No. 3, that says: "What is your best estimate for the average mean size and bytes for the documents that are listed in this document as having been processed?"

A. That is my question there in No. 3 at the top, yes.

Q. Okay. And below that, is that Dwight Groff's

answer where he says: "There is no way to identify the mean doc size for documents processed during this time since the file size was not something that was tracked in our reporting model. We also think that this would provide very little value since documents can vary in size from less than 1 kilobyte -- or 1k to 20m."

Do you see that?

A. Yes. And that is Mr. Groff's response to my question.

Q. So Mr. Groff is saying there is no average mean size, right?

A. Well, he's saying there's no way to identify it, and I understood this to mean to identify based on actual data records. So I understood his response in the spreadsheets to be his best guess as to what -- what that average would be.

Q. So all of your calculations and graphs are based on Groff's best guess?

MR. SCOTT: Objection.

A. I would not say that.

Q. Well, you just said it's his best guess, right?

MR. SCOTT: Objection.

A. Well, your statement was related to all of the

charts in my report, and that is not the case.

Q. No. The charts that we've been talking about that relate -- that use the average mean size.

A. Let me refer back to my report to make sure I get the charts right. So we're talking about paragraph 55 in my report earlier where we mentioned Footnotes 28 and 29. Those calculations ultimately resulted in a number of 3.2 gigabytes per hour which then appears in Figure 10 of my report on page 31. And I believe it is only Figures 10 and 11 that depend upon that estimate of 150 kilobytes per hour.

With respect to -- and then with respect to Groff's -- Mr. Groff's other estimates, that is discussed in the previous paragraph, which is paragraph 54. And, for instance, we talked about 0.8 gigabytes per hour. That number is present in Figure 8, and Figure 9 depends upon that figure, page 29 in my report.

Q. And does that number also dependent upon what Groff says about average mean size?

A. Well, I only used it in Figure 8 and 9. I only used the values that were provided in the attachment to the e-mail from Mr. Groff that we had been discussing.

Q. What about the .8? Where'd that come from?

A. One moment.

Q. It's Figure 8.

A. Right. I see .8 there. Let me get the origin of it. So if you refer to Exhibit H to my report, there are -- the first six pages comprise Mr. Groff's response via e-mail; the subsequent three pages are the tabs of a spreadsheet that Mr. Groff attached in his reply; and the second of those tabs labeled March 2013 through March 2015, there is a value of 0.8 gigabytes per hour that Mr. Groff provided.

Q. And does that particular value also refer to average mean size -- the documents processed?

A. Well, I'd say it's related to it.

Q. How is it not the same thing?

A. Well, they're different quantities. One is the number of gigabytes processed per hour, and the other is the average file size. The average file size is measured in kilobytes, megabits, or gigabytes perhaps. And the value here is measured in gigabytes per hour. They're different types of number.

Q. They're different types of numbers, but don't the same qualifications apply where Groff says there's no way to identify the mean doc size for documents processed during this time since the file size was not something that was tracked in our reporting model?

A. Well, I agree that, you know, I interpret his response here to mean that there was no record of the sizes of the files that were processed through the ICCE platform.

Q. Well, do you also interpret his response when he says: "We also think that this would provide very little value." Do you agree with that?

MR. SCOTT: Objection.

A. Well, I wouldn't say I agree with it because -- because I wanted the answer to the question. Mr. Groff's impression of the value wasn't really something that -- that I cared about. I wanted to know what his best guess was for the average file size.

Q. So even though he says it provides very little value, you still think that, you know, the information that you came up with and the calculations you did and you present in your report are still good notwithstanding his qualifications on the values he's giving you?

A. Would you repeat that question?

MR. DOYLE: Could you repeat that, please.

(Question read)

A. Okay. So the word "value" is used in two different ways in that question, so I just want to clarify. I'm confident in the calculations in my report,

and Groff's -- Mr. Groff's comments about providing very little value were immaterial to me because I was just trying to come up with the best estimates that I could for the average file size.

Whether or not Mr. Groff thinks those values provide -- well, let's put it this way: Whether Mr. Groff thinks those quantities provide very little value, as he states here, to me is immaterial for my purposes. I don't understand what his internal valuation process is, but I -- I imagine it might be different from mine.

Q. But you used the numbers he provided you, correct?

A. I did use numbers that he provided me.

Q. And he's saying they have very little value, but you disagree with him?

MR. SCOTT: Objection. Asked and answered.

A. So Mr. Groff, I understand from this e-mail -- well, he makes a statement about the -- about the value of those quanties in terms of -- well, I don't know exactly what his determination of their value is, but I was using these numbers for a specific purpose. And for my needs, they -- they suited what I was looking for which is an estimate for the average file size.

Q. Okay. Isn't that estimate going to be flawed, though, if you're using numbers on average mean doc but the guy who's giving it to you says they're flawed and have little value?

MR. SCOTT: Objection.

A. Would you point me to where Mr. Groff uses the word "flawed" in this e-mail.

Q. Oh, you're right, he doesn't. He doesn't say "flawed." Let's go back and see what he says. He says: "There's no way to identify the mean doc size." Okay. We'll start with that. No way to identify it.

He then goes on to say: "The documents processed during this time, since the file size was not something that was tracked in our reporting model." That's the first thing he says about it. So they didn't track it.

Then he also says: "We also think that this would provide very little value since documents can vary in size from less than 1k to 20m." And your testimony is you don't really care?

MR. SCOTT: Objection.

A. So I don't know what Mr. Groff thought when he wrote this, but my impression is that when he says "provide very little value," I think he may be talking

about value to RELX's reporting model, and that's why it wouldn't provide value to them. As he mentions, the documents can vary in size, which I'm interpreting this to be less than 1 kilobyte in size to 20 megabytes.

Q. Well, he also says they weren't keeping track of this, right? But any file size, since the file size was not something that was tracked in our reporting model, right?

A. So I agree that RELX wasn't tracking file size, and my understanding is that Mr. Groff may have been trying to justify why they don't track it, because to them, it wouldn't provide them with -- or it would provide only very little value to them.

Q. But this is in response to your question what is your best estimate for the average mean size in files that are listed in this document as having been processed, right?

A. Well, Mr. Groff is replying to my question, but I think he's providing additional color here.

Q. What color is that that's additional?

A. Well, his initial response is essentially, well, we don't track this in our reporting model, so I can't tell you based on any data what this number would be. And in case you were wondering, the reason why we

don't track it -- and, you know, I'm speculating here as to what he thought -- but I think Mr. Groff may have been thinking let me also explain why we may not track this: Because the file size varies a lot and it wouldn't be useful for our reporting model.

Q. So you're adding all that in?

A. Well, I don't know exactly what Mr. Groff was thinking, but that's a potential explanation.

Q. Well, don't you think he might have been thinking we also think -- and this is his own words -- we also think that this would provide very little value. And by "this," he means the average mean size, right?

MR. SCOTT: Objection.

A. Well, you raise an interesting point. I think that may be why he uses the word "we" here. I think he may be referring to RELX in this case. That only makes me think more that he's referring to Informatica's -- I'm sorry -- RELX's opinion about why this would have value or no value or something in between for the reporting model.

Q. Well, doesn't he also follow that up with "since documents can vary in size from less than 1 kilobyte to 20 megabytes"? Meaning there's a wide range of documents, right, and they don't track the size of

documents on their system. Isn't that what he's saying?

A. I do think he's saying that they don't track the file size of documents processed. And he's also saying that -- I think he may be speculating a little bit, but he's saying that the documents can vary in size from less than 1 kilobyte to 20 megabytes which is --

Q. Well, what information do you have to believe that he's speculating?

A. Well, perhaps he isn't. Perhaps he was looking at something here, but he doesn't cite why he thinks this.

Q. Okay. But yet you still use the values he provided you on averages -- average file sizes, right, notwithstanding this information, correct?

A. Well, like I said earlier, I was looking for a number of different ways to populate the information in that core calculation document from Informatica, and this was one of the ways in which I sought to do that.

Q. You had no concerns about it?

A. Well, I'm not saying that Mr. Groff's estimates are definitely correct. I was just looking for a variety of ways to populate the document, and then picked the ways that were the most favorable to Informatica, and found that even in that case, at most,

this document calculated 12 cores.

Q. That was nice of you. So when you used his average means sizes, you used them because that was most favorable to Informatica?

A. Well, I'm talking about all of the information that went into the --

Q. I'm not asking about that. I'm asking about the average mean doc size.

A. Well, it sounds like you were asking about the manner in which I provided information which was the most favorable to Informatica. And that information is the sum of information that was included in -- in the -- hold on one moment -- in the Informatica Sizing Model for RELX.

Q. Okay. Let's look at the next question. You asked: "To your knowledge, were there any periods of time between May 2010 to November 2017 where LexisNexis was processing documents with an average size nontrivially larger than the mean you estimated above?"

Answer: "No. See Item 3 above. Throughout all processing in the platform since it went live, the document sizes vary vastly across multiple content streams on a daily basis." Do you see that?

A. I do see that in this document.

Q. But yet notwithstanding that, in the next -- in the next -- his next answer, he gives a best guess estimate of 150 kilobytes, right?

A. Yes.

Q. And you used that -- that number, correct?

A. In one of the tables that we talked about earlier, yes, I did use the 150 kilobyte number.

Q. Was it one or two?

A. Let me refer back. So the 150 kilobyte number was used in paragraph 55 to determine the value of 3.2 gigabytes per hour, and then that was inputted into the Informatica Sizing Model for RELX as depicted in Figure 10 in my report, and then that resulted in the output for that document in Figure 11 in my report.

Q. Do you see where Mr. Groff says: "My best guess estimate would be an average file size of 150k, but I have no way to verify this"? Do you see that?

A. I do see that.

Q. Did you ever try to verify it?

A. Well, according to Mr. Groff, there is no data to verify it, and he states here that there's no way to verify. It seemed like there wasn't any way to verify it, so I didn't pursue that particular avenue.

Q. And then he follows with: "With this document

mean size assumption, we would have processed," and then there's -- I can't -- what is that number there? I can't even figure out what that is?

A. I think it's trillion to start.

Q. -- "trillions of bytes of data from 11/2012 to 11/2017."

Did you ever use that number, the total bytes of data that were processed -- or the guesstimate, I should say, on the total bytes of data that was processed from 11/2012 to 11/2017?

A. Well, I don't think I used that specific number, but I imagine you could probably come up with that number through some of the arithmetic in Footnotes 20 and 29 if you wanted to.

Q. Did you ever use that number for anything else? Any other charts? Analysis?

A. The 146 trillion number?

Q. Yeah.

A. I mean, I don't think I used it anywhere.

Q. Did you produce the spreadsheets used to create Figures 12, 13, 14, and 15?

A. I don't think these were produced, not the raw data that was used to make these.

MR. DOYLE: All right, Counsel, we request

that be produced immediately.

MR. SCOTT: Before or after you send me the stuff I requested last week?

MR. DOYLE: And we'll keep this deposition open too. What we need is the underlying data for all these figures --

MR. SCOTT: This deposition is not going to remain --

MR. DOYLE: -- as well as the formula used for waiting in Figures 14 and 15.

MR. SCOTT: I ask that you put your requests in writing, Scott.

MR. DOYLE: I will do that too, but I'm putting it on the record.

MR. SCOTT: That's fine.

MR. DOYLE: We also want to ask for the actual spreadsheets that the witness used to create Figures 12, 13, 14, and 15.

MR. SCOTT: I'll give you the same answer you gave me last week: We'll see if that becomes necessary. Is that what you told me last week when I asked you for the same information?

MR. DOYLE: You know what? I'm going back to questions.

Q. Sir, did you do any independent analysis to determine file size?

A. When you say "file size," do you mean --

Q. The average -- the average file size. The same thing we've been talking about for 20 minutes.

A. I think with respect to average file size, I asked Mr. Groff if there was data available for that, he said there wasn't, so I used his estimate in a couple of the paragraphs in my report that we've been discussing previously.

Q. Did you ask anyone else to -- information about what the average file size would be?

A. Well, I don't know if he did or not, but he doesn't mention anything like that in his responses to me in this e-mail.

Q. Well, did you?

A. No. I only asked Mr. Groff whether or not there was data, and then I asked for his own estimates.

Q. Why not?

A. Well, Mr. Groff was provided to me as someone who was knowledgeable about the Informatica system as it was used in the ICCE platform, and so I figured his estimate was probably a decent one. And like I said earlier, I was trying to find various ways to populate

this calculation document for calculating the number of cores that might have been appropriate for RELX's system.

Q. So the file size number you used is pure speculation, not based on science. It's not scientific at all, is it?

MR. SCOTT: Objection.

A. Well, Mr. Groff says it's his best guess estimate. And this is only one of the numbers that I used in order to populate that document. Some of the other numbers -- well, another number was the number that was actually in the document as produced. I was trying to, like I said, come up with a number of different ways in order to populate this document and then come to a conclusion based on those inputs that provided the highest number of CPU cores that the document would calculate. And, in fact, the estimates from Mr. Groff resulted in a higher number of CPU cores calculated than the data as originally populated in the spreadsheet.

Q. What is your background in capacity planning?

A. Can you define "capacity planning"?

Q. Have you ever heard of "capacity planning" in terms of developing in computer software?

A. I have heard the term "capacity planning."

Q. And what's your understanding of it?

A. In general, my understanding of capacity planning is it's -- it's the way you might plan for how many resources you would need in a system in order to execute certain software that was doing certain tasks.

Q. And did you do that? Do you have a background in doing that?

A. I specifically haven't been hired by any company to do their planning capacity.

Q. Do you have any other background in capacity planning?

A. Well, I do have a background in computer science, but specifically for the question of capacity planning, I don't think I've been retained, with the exception of this case, to examine issues related to capacity planning.

Q. Well, do you know when you're doing capacity planning that you need to determine and consider peak loads as well as the average loads?

A. I think depending on the system, that would probably make sense in certain circumstances.

Q. Did you consider peak loads in your analysis here of the Informatica software on the ICCE platform?

A. Well --

MR. SCOTT: Objection.

A. -- in this particular case, I was using the document produced by Informatica that was -- that appeared to be related to planning for the number of cores that a system like the RELX system would be using.

Q. Did you ever look at actual data relating to peak loads?

A. Well, by peak loads, if we're talking about CPU utilization, that was present in Expert Discovery 8 where it had average CPU utilization for individual servers for every hour of the time frame at issue.

Q. Other than average utilization, did you look at anything else to consider peak loads?

A. Well, in the context of capacity planning, I was looking specifically at the document produced by Informatica that seemed to calculate the number of cores that would be appropriate for a system. It would make sense to me that such a document would take into account various issues related to capacity planning, and peak load may have been one of them.

Q. Is that your understanding: that looking at average utilization is the same as looking at peak loads?

A. I wouldn't say it's the same. It's one way to get some information about certain resources, in this case, CPU cores that might be utilized more or less

heavily at certain times. There are probably other things you might consider when looking for peak loads for a different definition of "load."

Q. Are there any better metrics to use?

A. Well, I think better in this context depends on the question you're trying to answer. You might think that, for instance, if peak loads means, well, we're using a lot of the memory of our computers during peak loads, maybe we'd want to look at memory capacity; or maybe it means the number of tmp files I created when there's a lot of processing going on, so maybe we'd want to look at that. Maybe it's network bandwidth. There are a number of different things we could look at.

In this particular case, CPU utilization seemed to be one of the most relevant ones because the licenses are distributed on the basis of cores and not other resources that the computer system would have at its disposal.

Q. What about peak number of documents?

A. If documents are the way you're defining peak load, you could look at the -- you know, when certain numbers of documents were being processed that were higher than others.

Q. And did you look at that in this case?

A. Well, specifically the opinions I was offering -- well, one of them was related to the degree to which RELX may have benefited from having its ICCE platform have access to more cores rather than fewer. So I was focused on the actual processing of the documents as reflected in the CPU utilization.

Q. So you never looked at the peak number of documents or peak workloads?

A. Well, I did look at number of documents processed over time. I don't think I specifically looked for instances where the numbers of documents processed on a certain day might be higher or lower except with respect to increasing number of documents that were processed by the system over time.

MR. DOYLE: Why don't we take a five-minute break?

MR. SCOTT: Sure, Scott.

THE VIDEOGRAPHER: The time is 4:53 p.m. We are off the record.

(Recess)

THE VIDEOGRAPHER: The time is 5:12 p.m. We are on the record.

Q. Mr. Rucinski, before we broke, we were talking

about capacity planning. Do you have any training in capacity planning?

A. Specifically with respect to capacity planning, I don't have any training besides, generally, the computer science degree that I have.

Q. Thank you.

MR. DOYLE: Okay. I'd like to mark as Exhibit -- what are we on now?

(Exhibit No. 4, Expert Discovery 000002 Excel spreadsheet marked for identification)

Q. Now, on paragraph 68 of your report, you discuss ICCE files process report; is that right?

A. Give me one moment. Yes. In paragraph 68, I do see my discussion about the ICCE files process report.

Q. Okay. So if you look at Exhibit 4, and the first page of Exhibit 4, is this Expert Discovery 000002?

A. This does appear to be an excerpt from that document, yes.

Q. And this shows the number of documents that are processed by ICCE?

A. This, I believe, does show the number of documents processed by the ICCE platform over time.

Q. And does it show it on a per day basis?

A. This -- yes, this does appear to show the number of documents processed each day.

Q. Who prepared this spreadsheet?

A. The spreadsheet was produced by RELX. I don't know exactly who prepared it specifically.

Q. Did Mr. Groff prepare it?

A. He very well may have, but I don't know specifically whether he did.

Q. Was it -- was this document prepared for this litigation?

A. I believe it was since it appears the Bates number's Expert Discovery 2.

Q. And do you know if it was prepared specifically for your report?

A. I don't know why it was produced. I did use it in my report.

Q. You didn't use it in your report?

A. No, I did use it in my report, but I don't know if was produced specifically for that purpose.

Q. Okay. Did you ask for this data?

A. I did want to see the total number of files processed through the ICCE platform on certain days, and this document seemed to show that.

Q. Okay. So if we go to the next page, do you

see where it says, "On the date of April 10, 2015"?

A. I do.

Q. Okay. And the highlighted line shows the -- that ICCE processed a total of 2,025,684 documents?

A. Yes, I do see that line.

Q. And so you're familiar with this particular sheet of the spreadsheet?

A. I am familiar with this tab of the spreadsheet, yes.

Q. And do you believe this data is accurate?

A. This was the data provided for me to this topic, so I do believe this data is accurate.

Q. And did you perform any independent analysis to determine its accuracy?

A. Not that I recall. I think I was wondering about how many documents went through the system, and this was a document that seemed to provide that.

Q. Why didn't you?

A. Well, like the other documents that were provided, this seemed to answer the question, and so I had no reason to doubt its accuracy.

Q. Did you ask how it was created?

A. I believe in conversations with Mr. Groff, well, he essentially said they had records to show that,

I think, and my understanding is that these were those records.

Q. Mr. Groff said what? They had records to show what?

A. To show the number of files that went through the ICCE platform on various days.

Q. Did you ask how it was created?

A. I wasn't sure exactly how this chart was created. It was provided to me by RELX.

Q. How was the data for the spreadsheet collected?

A. Well, I don't know exactly how it was collected, but it seemed like RELX had records of the number of files that went through the platform on various days, and this is what those records showed.

Q. Where did it come from, the data?

A. Well, it came from RELX's records, but I don't know specifically where it came from from RELX.

Q. Was it filtered in any way?

A. I believe it may have been filtered based on uniqueness.

Q. Based on what? I'm sorry.

A. The uniqueness of files.

This is something I wasn't aware of as I was

writing the report, but my understanding now is that I believe this document -- I'd want to verify this elsewhere -- but my understanding is that this document may include only unique files that were processed through the system and not just the total number.

Q. And what do you mean by "unique files"?

A. Well, if you have two separate collections of data comprising a file, you could say that you have two separate files, but if the data is the same, then they are in some sense duplicates, and so if you wanted to count the number of files that had unique data, then those two files that are the same would count as only one file.

Q. So in this chart that we're looking at, does this show the number of unique files or the number of total files which would include non-unique files or duplicates?

A. I think -- this is something I'd want to verify, but I think this might show unique files. I'm not sure whether it's unique or just total files in general.

Q. So you don't know?

A. I'm not sure, as I sit here right now.

Q. Have you asked anybody from RELX?

A. Well, I discussed -- I was only made aware of the potential for some of this data being filtered for unique files as of this morning, so I didn't have an opportunity to talk to RELX about it.

Q. Who told you this morning that it may be filtered -- did you say only for unique files?

A. Well, I think what I said is that this may -- well, only -- if I said only by unique files, what I mean is that this may include only account of unique files rather than total files, but I'm not certain about that. And that information came from counsel.

Q. Did you talk to Mr. Groff this morning?

A. No, I didn't.

Q. Did you talk to Mr. Groff yesterday?

A. I did for ten minutes, like I mentioned earlier.

Q. Did he refer to this issue?

A. He did not mention it.

Q. So you learned about this from counsel?

A. I did.

Q. What did counsel tell you?

A. Counsel said there may be a discrepancy in the number of files that were reported as passing through the ICCE platform with respect to total

files versus unique files.

Q. What would you like to depend on, unique files or total files, for your analysis and opinions?

A. Well, there's a couple of different questions I want to answer. Let me think for a moment. So regarding the question of how many files that ended up on the Lexis Advance or Lexis.com products versus the number of files that went through the ICCE platform, I think what I'd want there is unique files, because in that case, what we care about is the percentage of files that were actually provided to RELX customers versus -- or rather, the percentage of files provided to RELX customers that went through the ICCE platform.

So whether or not there were duplicates of those files on the platform I don't think really matters because we're counting based on what a customer would actually see, and they don't see duplicates.

With respect to the question of benefit, I do use file totals through the -- or total number of files processed through the ICCE platform to look backwards for time periods during which there is no utilization data available. And I think for that question, it might depend on, to the degree there are duplicates, where those duplicate files arise, because in that case,

we're -- the question at issue is the number of files that were touched by Informatica software and not in general by the ICCE platform.

So I think in that case, unique files might be the safer bet unless I had more information about why there were duplicates and at what points in the process those duplicates were created.

Q. Do you plan on supplementing your report or your opinions based on answers to these questions?

A. I'm not sure at the moment. I don't have any explicit plan to right now, but I imagine that might happen depending on how the case turns out.

Q. What paragraphs may you have to change or that might be incorrect at this point?

MR. SCOTT: Objection.

A. So I think the paragraphs that would be subject to potential revision would be paragraphs 34, 68, and 69, including the charts in those paragraphs.

Q. So that was 34. What's the next one?

A. 68.

Q. Okay. And what else?

A. And as I sit here right now, I think the only other one would be 69.

Q. Do you see where -- in the last sentence where

it says: "The maximum percentage of documents available on the Lexis Advance and Lexis.com products that were processed by the ICCE platform according to the ICCE processed documents percentage report is 15.15 percent on September 6, 2017"? Do you see that?

A. I think it says 15.05 percent.

Q. Thank you for that clarification. Do you think that may have to be modified?

A. Yeah, I may need to modify that sentence either with respect to the number or the way in which it's described.

Q. And what about Figure 5? May that have to be modified?

A. It's possible I would modify that too.

Q. Okay. And I think the next paragraph you said is paragraph 69?

A. I believe I said 68 and 69.

Q. Okay. Do you see in the last sentence where you say: "And the hourly weighted average CPU core, equivalents during the early time period was therefore almost certainly not greater than either the hourly weighted average CPU core equivalents during the middle time period or late time period"? Do you see that?

A. Are you looking at paragraph 68?

Q. Yeah, 68, the last sentence.
A. Are you looking at the sentence that starts with "since the instructions"?
Q. Yeah.
A. Okay.
Q. Why don't we just -- I'll revise my question to be with respect to that entire sentence, might that have to be modified?
A. I think it's possible but unlikely. I don't really expect this to change unless there's some vast difference after looking at the -- the uniqueness of the files.
Q. And, then, what may change with respect to paragraph 69?
A. Well, I may change the graph at the end.
Q. Figure 16?
A. Yes, Figure 16.
Q. Anywhere else in your report that you may want to change?
A. As I sit here right now, that's all I can think of.
Q. Do you know who's going to go through all this data and determine what's unique and what's not unique?
A. Well, I haven't asked anyone to yet, so I

don't know who ultimately would do that. I imagine I'll go through it myself, but I'll have to talk to someone to -- probably from RELX to make that determination.

Q. And who would that likely be?

A. I'll probably talk to Mr. Groff first and see what he says.

Q. All right. Back to Exhibit 4, if we could. So we were showing that for April 10, 2015, the highlighted line shows that ICCE processed 2,025,684 documents, right?

A. Yes, that is the number of this spreadsheet.

Q. Okay. If you go to the next page, is this printed from Expert Discovery 000010?

A. Yes, I believe it is, at least given the title at the top.

Q. And if you look at this, this chart -- this chart shows the number of documents as a total by day; is that correct?

A. I'll look at it for one moment. Yes, I believe that's the case.

Q. So if you look, for example, at that date of April 10, 2015, which is the same date we were looking on the previous page, do you see that there are 44,796,536

documents in the repository from ICCE?

A. That's the number that's on the spreadsheet, yes.

Q. Right. And that's the total documents for that day; is that correct -- for ICCE?

A. Yes, that's under the ICCE total document count header, so I do think that's the total document count for ICCE that day.

Q. And isn't it true that if you wanted to see the number of documents that were added from the previous day, or 4/9/2015, you would just subtract the previous day's total?

A. Give me one moment. I think that's generally the case. The reason I'm hesitating about this for a moment is because if that were the case, you would expect the document counts here to be monotonically increasing, which is to say that they wouldn't decrease on a particular day.

For the Product Total Document Counts column, which I understand is analogous to the ICCE Total Document Count column in Row 455, the number on that day, April 6, is less than the day before it. So I think this may still be cumulative, as you suggest, but there may be certain additions or deletions that are accounting for

some of the discrepancies here.

Q. And what would account for the discrepancy in going from a higher number to a lower number. Is that what you're saying?

A. Right. So there could have been documents that were removed from the system. That's what I'm saying.

Q. Okay. So let's look at the difference between April 9, 2015, and April 10, 2015. Do you see that --

A. I do.

Q. -- on the ICCE total document count?

A. I do.

Q. So if you subtract the value that's provided on April 9 from the value provided on April 10, that will show you how much you add -- were added in that particular day; is that correct?

MR. SCOTT: Objection.

A. I believe it will show you the total count, so that would account for, potentially, additions and deletions, but you could see the difference that were -- that were processed through the system on the two days.

Q. Okay. Would you do that calculation? Maybe Mr. Scott can give you his famous calculator again.

A. I think it's 312,964.

Q. Why don't you go ahead and do it just to be sure.

A. Sure.

Q. Be accurate.

A. Okay. So I had -- ah, okay. I transposed the numbers. So -- there's a call from potential scam. Should I answer --

MR. SCOTT: No, we don't need to talk to them. That's why I have that on there.

A. Okay. If I entered this right, it's 181,949. Does that match what you have, or did I enter that --

Q. I'm sorry. What do you have?

A. 181,949.

Q. No, that's not what I get.

A. Let me -- let me try that one more time. That's not what I had either.

Q. You were closer the first time. You were real close.

A. So we're talking about -- oh, I'm sorry. The 10th minus the 9th. That is -- that is my problem. Okay. So that looks like 311,964. Does that match what you got?

Q. Yes.

A. Okay.

Q. So is it fair to say that on April 10, 2015, 311,964 documents were added --

MR. SCOTT: Objection. Asked and --

Q. -- through the ICCE?

A. That's my impression. Well, I think this document is showing the total number from the ICCE platform on the various dates, so, yeah, I think the difference would show the difference in totals between the two days.

Q. Okay. Which, in this case, is 311,964, right?

A. If that's the same number I just entered on the phone --

Q. Yeah.

A. -- yes.

Q. Okay. Now, did you compare that number, which is set forth in Expert Discovery 000010, to the number for that same day in Expert Discovery 00002, which is on the previous page?

A. No, I don't think I compared those two numbers.

Q. Well, if you look on the previous page, which is the 000002 document, it shows a total number of documents processed on the ICCE platform of 2,025,684, whereas on the 000010 document, it shows 311,964. Can

you explain the difference?

A. As I sit here right now, I don't know why the discrepancy is present in these documents.

Q. Shouldn't they be the same?

A. Well, it depends what these documents are actually showing.

Q. You don't know what the documents show?

A. Well, like I said earlier, I think one of them might be related to uniqueness of documents and the other one may not be.

Q. But you don't know?

A. As I sit here right now, I'm not sure. I'd have to talk to the folks at RELX.

Q. You had mentioned before that you were surprised that -- that it looked like documents may actually go down from one day to another day; is that right?

MR. SCOTT: Objection. Mischaracterization.

A. I noted specifically on -- this is page 3 of Exhibit 4, that there was -- that there was a day where the total -- the product total document count in the document was less than the previous day.

Q. So I'm sorry. Where are you looking at? Which --

A. So I'm looking at Row 455 and Row 454 under the Column C.

Q. The Product Total Document Count?

A. Yes, that column.

Q. And so it shows less documents per product total on 4/6/2015 from 4/5/2015?

A. It does, yes.

Q. Does the ICCE platform delete documents?

A. Well, this is respect -- with respect to the product total document count. So my understanding is that this number is related to the number of documents that are available through the Lexis.com and Lexis Advance products. So it seems to be that they sometimes do remove documents from the system, according to these numbers.

Q. Would that affect your analysis if you knew that documents were actually removed from the system from day to day?

A. Well, so long as the two columns here are -- make sense in terms of ratio, which I think they still do, I don't think that would affect it. It may change the way in which I describe that number, but overall, that number is a percentage of documents processed through the ICCE component of the platform that actually

make it to the -- to the customer at the end.

Q. Are you surprised that the totals went down from one day to the next day?

A. I think I actually did notice this prior, but it made sense to me that perhaps sometimes documents are removed from the system.

Q. How are you going to determine this discrepancy between the number of documents added on April 10, 2015, and the ICCE total document count on the 00010 [sic] spreadsheet as compared to the different number provided on the 000002 spreadsheet?

MR. SCOTT: Objection. Asked and answered.

A. So first of all, I'll talk to someone from RELX, which will probably start out being Mr. Groff; and then we'll probably have to determine the underlying data that was used to generate these documents and account for the discrepancy in some way.

Q. I mean, that's a huge discrepancy, I mean, between what's shown on 00002 [sic] of 2,025,684 documents, and what's shown on Expert Discovery 000010 which is much more, 2,025,684, right? The factor of 10, approximately?

A. Well, that may be easily explained by the uniqueness issue that I mentioned earlier. I'm not sure.

I'll have to look into it.

Q. Do you know if these errors change your opinions in --

MR. SCOTT: Objection.

Q. -- any way?

MR. DOYLE: Can I get my whole statement out before you object?

MR. SCOTT: I'll try.

MR. DOYLE: Please do.

MR. SCOTT: I am trying, Scott. Sometimes it's difficult, as you know.

Q. Did you get the question, sir?

A. Would you repeat it?

Q. Do the errors we just identified change your opinion in any way?

A. So this discrepancy relates to specific portions of my opinions. I mentioned that I may want to revise portions of them, but I don't think overall they will affect my -- my opinions.

Q. Now, we've given you examples of some errors. How are you going to become comfortable and satisfied that there aren't other errors through all this information?

MR. SCOTT: Objection.

A. Well, one way in which I'm uncomfortable is that I presume that if Ms. Frederiksen-Cross has found such errors, that they'll be brought to my attention either at deposition or through some other means. But I will have conversations with Mr. Groff to talk about the origin of data. I don't think the specific issue with respect to uniqueness of the files on these two sheets affects, for instance, CPU utilization because they're just completely different types of information, but it's worth talking to Mr. Groff just the same.

Q. Well, we don't -- we don't have the underlying data that actually exists at RELX. Do you understand that?

A. Well, I think the numbers in all of these spreadsheets we looked at come from data at RELX. I think the question at issue here is what that data represents.

Q. Right. But what I'm saying is because we don't have the underlying data at RELX, how are we going to determine whether or not the underlying data is actually correct and accurate?

MR. SCOTT: Objection. Calls for speculation.

A. I think we -- the underlying data, I think, is produced in documents. I don't think we have to go to

RELX to look at the same data. They can send it to us in a document.

Q. Oh, so the data that populates these spreadsheets is actually in other documents?

MR. SCOTT: Objection.

A. Well, I don't think it exists only in this document that I have here produced for this litigation. I believe it was created based on data that RELX has that they store in their ordinary course of business.

Q. And have you seen that document as they store it in the ordinary course of business?

A. I believe I've only seen the documents that are produced in this litigation.

Q. That's not my question. I just asked you have you seen the documents that RELX keeps in the ordinary course of business --

MR. SCOTT: Objection.

Q. -- which contain this data?

A. Well, to the extent those documents are produced in this litigation, then I have seen them. But to the extent they're not produced in this litigation, then I haven't seen them.

Q. Well, you've testified that you received, in many instances, the spreadsheets from Mr. Groff, correct?

A. Well, I don't think I received them directly from Mr. Groff, but from RELX. Yes, I received spreadsheets from RELX.

Q. And is it true that -- did you investigate any of the data that populates the spreadsheets in any other document that was used to filter and put the data from the underlying documents into this spreadsheet?

MR. SCOTT: Objection.

A. Well, I relied on the data as it was produced and looked for any reasons why it may be -- why it may contain inaccuracies. And in this case, I seem to have missed one of those instances.

Q. You think there might be more instances?

A. It's possible, but I don't think that would be the case for the utilization data, for instance.

Q. Why is that? Did you look at every single piece of data, and did you verify every single piece of data for utilization?

A. Well, I looked through the data as produced and actually asked some questions to Mr. Groff about the date, certain time frames that were missing data, for instance, and he provided explanations that seemed reasonable and helped to assuage any concerns that I had about it.

Q. Can you give me an example of that?

A. One moment. So one example is in Exhibit H to my report, looking at page 2. At No. 3 towards the middle of this page, I state in this e-mail to Mr. Groff: "Please refer to RELX v Informatica document production ICCEUsage.xlxs, which was produced in the Informatica litigation." And then at No. 2 below that identifying a particular tab with a particular range of rows that have a different machine name.

And so I asked a question to Mr. Groff about what was actually being run on this server during that time. Mr. Groff replies: "During this time, psc33817 went completely down and was not available for any processing. When this server was brought back online, the monitoring tool for collecting metrics for the server was not configured the same way, causing the label to be different."

So this explained one of the discrepancies that I noticed in the data produced.

MR. DOYLE: Court reporter, please mark as Exhibit 5 a very large sheet with a lot of information on it relating to batch start times and end times. And it's also the data that is from September 25, 2013, identified as RELX 044494.xls.

(Exhibit No. 5, two-page spreadsheet marked for identification)

MR. SCOTT: This is 5, you said?

MR. DOYLE: Just for the record, these two pages actually go side by side. So the correct way to look at it is to put it side by side.

Q. Sir, have you seen these document -- this document before, or any others like it for other days?

A. This does look familiar, but I don't know if I've viewed this specific document.

Is there a question pending?

Q. Yeah. Have you seen this document before?

A. I think I said that I don't remember if I've seen this specific document, though it does look familiar.

Q. So you've seen other documents like it?

A. I believe so, yes.

Q. And what does it tell you?

A. Well, it looks like for a specific day -- maybe part of a specific day, it looks like it shows -- it looks like it shows things like something called delivery bundle ID, DPSi, content type, ICCE conv batch name, start and end times, a number of xml documents for a number of other header columns, an ICCE conversion

batch name, at least that's what I assume "conv" stands for. And then there are a number of columns related to FPD which appear to be unpopulated, and then unpopulated columns related to delivery complete or delivery error.

Q. Did you cite this document in your report?

A. I might have. What's the Bates number of it?

Q. RELX 044494.

A. Yes, I did cite it in my report.

Q. Okay. Why might you want to look at this document?

A. Well, I think there were documents like this that showed for specific days, for instance, batch start times and end times, but I never saw a document that -- that spanned a time frame long enough that I felt was relevant to the question of what the processing time was for the different documents going through the ICCE platform.

Q. Do you recall earlier today that you had not seen any documents that show processing time?

MR. SCOTT: Objection.

A. I may have testified to that effect, and if I did, I forgot that these documents for individual potential portions of dates existed. I was thinking

about documents that showed batch processing times across the relevant time period.

Q. Did you ask for more of these documents?

A. I did ask for documents that showed this across a time period.

Q. What time period?

A. I believe it was for the entirety that the ICCE platform was processing documents.

Q. Which would be what?

A. That would be approximately November 2012 through November 2017.

Q. And what was the response?

A. I don't think I got a response with those documents, and so I presumed that they didn't exist for that time period.

Q. So because you didn't receive the documents, you presumed they don't exist?

A. Well, I think I asked for them and never received them, so that was -- that was my conclusion.

Q. Who did you ask?

A. I think I spoke to Mr. Groff and also counsel about it.

THE VIDEOGRAPHER: There's seven minutes left of video disc.

MR. DOYLE: Want to take a break? Do you have time?

THE VIDEOGRAPHER: The time is -- the time is 5:55 p.m. We are off the record. This is the end of Disc 4.

(Recess)

THE VIDEOGRAPHER: The time is 6:29 p.m. This is the beginning of Disc 5. We are on the record.

Q. Mr. Rucinski, you reviewed some of the copyright registration certificates of Informatica in this case?

A. I did review some of those copyright registration certificates, yes.

Q. And do you understand that a copyright registration certificate is evidence of the validity of the copyright?

MR. SCOTT: Objection. Calls for a legal conclusion.

A. While I'm not here to offer legal opinions, I do understand that the copyright registration is something that's -- that's submitted when applying for a copyright at the copyright office.

Q. Why did you review the copyright registration in this case?

A. Counsel asked me to review what was submitted for the copyrights assigned to Informatica and characterize what was included in those deposits.

Q. Do you have any opinions as to whether or not how those relate to this -- the copyright registrations relate to the copyright infringement allegations in this case?

A. Well, my opinions are stated in my report, and none of them relate to legal matters. My opinions are related, with respect to the copyright deposits, in terms of what was actually in the deposits, specifically with respect to whether the deposits contained source code that is human readable or executable code that's not.

Q. And do you know whether a copyright registration with source code as a deposit would be effective in a copyright infringement case based on the copying of executable code?

MR. SCOTT: Objection. Calls for a legal conclusion.

A. That sounds like a legal opinion to me, and I think that's outside the scope of my opinions relating to computer science in this case.

Q. Have you ever reviewed copyright registrations before?

A. Yes, I have reviewed copyright registrations before.

Q. In what context?

A. There was a litigation I worked on that related to copyright infringement where we were working for the -- one moment. This was a while ago. We were working for the plaintiffs in this case, and there was a question about copyright ability of certain code as well as website material in the case, and so I reviewed copyright registrations for -- for those copyrights.

Q. You said counsel asked you to review the copyright registrations?

A. Counsel did ask me to do that, yes.

Q. And why was that?

MR. SCOTT: Objection. Calls for speculation.

A. My understanding is that counsel wanted me to examine the registrations and characterize them in some way. I don't know exactly why they wanted me to look at the specific registrations.

Q. The copyright registration certificate for the B2B data transformation 9.6.1 describes the copyrighted material as a computer program, right?

A. Let me refer to my report. Which specific Informatica component were you talking about?

Q. 9.6.1. The B2B data transformation 9.6.1.
A. And would you repeat that question?
Q. Sure.
MR. DOYLE: Can you repeat it back, please.
(Question read)
Q. I'll read the question again.
Is it your understanding that the copyright -- copyright registration certificate for B2B Data Transformation 9.6.1 describes the copyrighted material as a, quote, computer program, end quote?
A. I think it lists computer file as the type. At least that's what I wrote in my report.
Q. That the copyright registration certificate describes the copyrighted material as a computer file, is that your --
A. That's my recollection. That's what I recorded in my report. So with respect to copyright registration, I think that's what it says.
Q. Okay. And did you also look at the copyright office's catalog entry for B2B Data Transformation 9.6.1?
A. If you're talking about the copyright deposit, yes, I think we're talking about the same thing.
Q. And what was that -- what did that describe the computer copyrighted material as?

A. As I sit here right now, I don't remember what the cover page described it as. I was focused on the computer source code that was attached.

Q. And do you understand that the copyright registration certificate for B2B Data Exchange 9.6.1 describes the copyrighted material as a, quote, computer program, end quote?

MR. SCOTT: Objection. Asked and answered.

A. And we're talking about data exchange now?

Q. Yes.

A. So I believe the copyright registration describes -- or lists, quote, computer file as the copyright material for that copyright. It may say something different on the cover page for the copyright deposit. I don't remember, as I sit here right now.

Q. And is it your understanding the copyright office's catalog entry for B2B Data Exchange 9.6.1 describes the copyrighted material as a, quote, computer file, end quote?

A. I believe the registration that's -- that's publicly available online has computer file listed as the -- as the type of work.

Q. Did you also review Informatica's copyright registration certificate for the Informatica

PowerCenter Grid Option?

A. I did review that registration, yes.

Q. And do you understand that the copyright registration certificate for PowerCenter Grid Option describes the copyrighted material as a, quote, computer program, end quote?

A. Well, I looked at the copyright registration information available online for the type of work which I believe listed computer file. So I don't think I understand what you mean by "the certificate" in this context and for the previous items.

(Cell phone)

MR. DOYLE: Can't stop a wife from calling.

MR. SCOTT: That's how you stay married. But you should answer it too.

MR. DOYLE: No, I know because it'll probably keep -- it'll keep coming and coming and it won't stop and the anger -- anger will build.

MR. SCOTT: Yeah, avoidance is not the answer.

Q. Sir, are you familiar with the difference between a single core CPU and multicore CPU?

A. Yes, I'm familiar with the distinction between a single core CPU and a multicore CPU.

Q. Okay. And what is your understanding of the

advantages of a multicore CPU over a single core CPU?

A. Well, in general, there are -- there are some tradeoffs and benefits to both. So single core CPUs generally have a higher clock speed for the one core just because there's more space on the actual chip to -- to provide that speed.

For multicores, supposing you have two cores for a single CPU, generally, the speed for each core is lower, but in exchange, what you get with a multicore CPU is you can have parallel processing of tasks when the computer program is run. So, for instance, if you have a computer program that is executing multiple threads to perform different tasks, you can execute each of those threads simultaneously, whereas in a single core case, you would have to wait for -- or kind of put one task on hold while the CPU deals with the -- with the other one and vice versa.

Q. Can a multicore CPU process more workloads at the same time as compared to a single core CPU?

A. In general, yeah, multicore CPUs can process multiple tasks at the same time or multiple threads. And if the workflows for the Informatica software for certain versions were programmed in that way, then I imagine a workflow, if it were part of the single thread, could

execute on one CPU while another workflow on another thread could execute on another CPU core.

Q. Why would you want to execute threads in parallel?

A. Well, there could be a few reasons why you might hypothetically want to execute threads in parallel. One reason might be that you wanted to perform two different tasks at the same time while using the same portions of memory. The threads are in the same address space, so that would be possible.

It could be that you have a single process that has components which are not dependent upon one another, and so you can execute those two components at the same time and then combine the results as opposed to processing it one at a time.

Q. Are there advantages to using all the available cores on a server?

A. You're asking that question in general?

Q. Well, with respect to a multicore CPU, are there any advantages to using all of the cores on that server?

A. Do you mean all the cores versus fewer of them?

Q. Sure.

A. Well, if the program you were running could make productive use of all the cores as opposed to just fewer of them, I think, in theory, such a program would be able to be more productive over -- over a certain time period.

Q. Can running tasks in parallel reduce the total time required to process multiple tasks as opposed to running one at a time?

A. Well, in general, it would depend on the speeds of the -- of the cores on the CPU versus the hypothetical single core that we're discussing. But there are circumstances under which if you have a program that can take advantage of multiple cores, because of the efficiency you would get by executing different components of a process in parallel versus in serial, you could, in theory, get a result quicker on a multicore processor than with a single core.

Q. Is Informatica software the kind of software that can use more than one CPU core at a time?

MR. SCOTT: Objection. Asked and answered.

A. As far as I know, there are specific circumstances under which that might occur for specific versions of Informatica software.

Q. And are you aware that the Informatica

software is the kind of software that can use multiple cores at the same time?

MR. SCOTT: Objection. Asked and answered.

A. Sorry. Was that different from the previous question?

Q. No. I meant to say is the Informatica software, could it use multiple cores for concurrent workflows?

A. I'm not entirely sure about that only because I'm thinking back to Ms. Frederiksen-Cross's report with her testing. I don't remember if she specified the number of workflows that were part of the test. I think she specified the number of files, but I don't recall as I sit here right now.

Q. Do you know where else you might be able to obtain that information other than Ms. Frederiksen-Cross's report?

A. It's possible there might be documentation that speaks to that question, but as I sit here right now, I'm not sure of the answer.

Q. If you wanted to determine how many cores were used during processing some hypothetical workload, how would you do the testing?

A. To be clear: Your question is about

processing a single workflow?

Q. Sure.

A. In this hypothetical test, are we talking about coming to a conclusion about how the Informatica software would operate in certain circumstances, or are we trying to generalize?

I'm just trying to get a feel for what the goal is that you --

Q. I was just asking in general what kind of testing would you do to determine how many cores were used during processing some hypothetical workload.

A. Well, if we cared about a specific workload and we just wanted to see for that one specific workload what would happen for a specific version of Informatica software on a particular hardware setup, I would imagine probably the way to do that would be to install the Informatica software on that hardware setup that we specify for this particular instance, identify the files that we care about processing in this one specific workflow, and then monitor over some period of time the degree to which the Informatica software in this specific case used the different cores that were available to it. And the overall conclusion of such a test would be limited to the specific setup that we had provided.

Q. What do you mean -- oh, the conclusions of the test would be limited to the particular setup that was -- you were testing? Is that what you mean?

A. So in this hypothetical example, my understanding is we're talking about seeing what would happen for a very specific workload if it were run through the Informatica software. So if want to see how it would work for the very specific workload, it makes sense to me that we would run that workload through the Informatica software for a specific version of the software and the specific version of hardware that we cared about, and then see -- see what happens.

Q. And what did you mean "see what happens"?

A. Well, we could monitor -- we could monitor the CPU cores and see when they were used for processing the workload that -- that we cared about for a specific test.

Q. How would you test to determine whether Informatica makes use of multiple cores concurrently?

A. You're asking in general or for a specific workload again?

Q. In general.

A. Well, if we wanted to generalize, we would probably have to, from an empirical standpoint, install

the software on a few different hardware configurations. And then if we wanted to generalize, we would have to come up with a number of different workflows with a number of different numbers of files in each of them, different types of files.

We would also -- we would have to consider which versions of the Informatica software we were using. We'd have to run it for some amount of time that we were comfortable with to generalize to -- to larger amounts of time. For instance, there -- you know, Mr. Groff has stated that there are instances where certain files got -- got stuck or workflows got corrupted, so we'd want to see how -- well, we'd want to test until we encountered situations like that so we could see how it might affect -- how the software was using the various cores that were available to it if we wanted to get to a general case.

There may be other considerations. As I sit here right now, those are the ones that I can think of.

Q. RELX used several servers in a grid; is that correct?

A. So the ICCE platform used a number of servers as part of the ICCE platform that included Informatica

software for part of the time period at issue in this case.

Q. Did it use it in a grid?

A. My understanding is that the servers were configured in a grid in order to use the ICCE platform as well as the Informatica software that was incorporated as part of it.

Q. What is a grid?

A. My understanding of a grid is it's a mechanism by which you can organize multiple servers in order to process certain tasks.

Q. Servers of different sizes or of the same size?

A. Well, given that my understanding is that the ICCE platform used servers with different configurations, it sounds like you could use servers in different configurations in the ICCE platform that incorporated the Informatica software for certain time periods, if that's what you mean by size. I mean, different configurations would at least be different between the servers.

Q. What type of servers were actually used in the RELX grid?

A. The specific servers, if my memory is correct, there were two types. One was a Dell R710 server, and

the other was a Dell R910 server.

Q. Okay. And how many cores do those servers have on them?

A. So my recollection, at least for certain periods of time, the Dell R710 servers had two sockets, and so two CPUs, each CPU had four cores, so that's a total of eight cores for the Dell R710 servers.

And then for the Dell R910 servers, my recollection is that there were four sockets available, and each of those sockets had a CPU that had eight cores, and so for those Dell R910 servers, there were 32 cores on the servers.

Q. And what processor speed were the cores?

A. I don't recall, as I sit here right now. I think it was around 2 gigahertz.

MR. DOYLE: I'd like the court reporter to mark as Exhibit 6 some e-mails, all of which -- well, a series of e-mails between various individuals including Chris Boytim, Gil Rosen, and others.

(Exhibit No. 6, e-mails marked for identification)

MR. SCOTT: What number is this one? 7? 6. Thank you.

Q. Have you seen these e-mails before?

A. I believe that I have.

Q. What are these e-mails in relation to?

A. It's been a long time since I've looked at them. Give me a moment. So in general, these e-mails, which are around the time frame of late 2014, appear to be related to trying to determine an appropriate number of licenses might be for RELX to have with respect to the Informatica software.

Q. Have you seen these e-mails before?

MR. SCOTT: Objection. Asked and answer.

A. I do think I've seen these e-mails before.

Q. And do you know whether or not these e-mails relate to the sizing model that you opine on in your report?

A. One moment. So I believe the document attached to the most recent e-mail in Exhibit 6 here, it's entitled "PowerCenter Size and Model, LexisNexisB2.xls," I believe that's the same document that I discuss in my report starting in paragraph 49.

Q. Okay. And then in paragraph 52 of your report, you state, quote, Informatica did not adhere to the calculations of CPU cores determined in Informatica Sizing Model for RELX, instead selling RELX far more CPU core licenses than the Informatica Sizing Model for RELX

calculated using the parameters below. Right?

A. I haven't verified it exactly, but that sounds like an accurate reading of the penultimate sentence of paragraph 52.

Q. What is the basis for that assertion?

A. Well, as I explain in my report, Informatica Sizing Model for RELX as produced contains calculations that suggest that, at most, six CPU cores would be appropriate, and then I did some additional analysis trying to see if there were reasonable other inputs to the document to increase that calculation to be closer to what RELX actually acquired in terms of licenses, and was able to come up with 12 as a maximum instead, which is less than the number of licenses that RELX did actually acquire.

Q. So this is in the 2014 time frame; is that correct -- these e-mails?

A. These e-mails, yes, they all seem to be from the late 2014 time frame.

Q. And are these e-mails about the sizing model that you refer to in your opinion, in your -- in your opinion?

A. In general, they appear to be about -- and there's maybe a dozen of them here. They appear to be

about planning for how many licenses RELX might want. And, then, in this last e-mail, there's a document attached which I believe is the same document I discussed in my report, so they seem related.

Q. And you believe that document to be the sizing model that you discussed in your report, the one that's attached to the e-mail sent on Friday, December 5, 2014, 6:59 a.m. from Michael Tomechak it Gil Rosen?

A. I believe that's one -- at least one place where this document shows up. I'm basing that based on the file name here.

Q. And that's the sizing model that you opine on in your report; is that correct?

MR. SCOTT: Objection. Asked and answered.

A. I believe it is based on the file name here.

Q. Yeah. And at this time, how many licenses did -- CPU licenses did RELX have at this time in 2014?

A. I'm trying to remember. As I sit here right now, I think it was 72.

Q. 72. And do you recall how many CPU cores the Informatica software was deployed on at that time in 2014?

A. Well, I wouldn't say the Informatica software was deployed on cores. It was deployed on servers. But

I believe at this time, this would have been after Nalin Mishra brought up the last three servers and before servers were removed. So I think at this time, there were seven servers which would mean that there were 104 cores available for the ICCE platform that incorporated the Informatica software.

Q. You just said "Nalin Mishra brought up the last three servers." What does that mean?

A. My recollection is that as an employee of Informatica, Mr. Mishra added servers to the ICCE platform that had the Informatica software deployed on them. I think that was also done by Mr. Mishra. That's my recollection from the -- some e-mails that I saw.

Q. Your recollection is that Mr. Mishra added servers to the network -- to the platform?

A. That is my recollection, yes.

Q. Are you going to give any -- do you plan to give any testimony at trial about whether or not the actions of Mr. Mishra relate to the copyright infringement in this case?

A. I plan to testify about the opinions that are in my reports. So to the extent those opinions relate to Mr. Mishra's actions, the answer would be yes. I'm not

sure exactly which opinions those would be, as I sit here right now.

Q. Did you review the service agreement between the parties?

A. There were a number of agreements. Would you clarify which agreement you're referring to?

Q. It's called the "Service Agreement."

MR. SCOTT: Objection.

A. I did review agreements between the two parties. I'm not sure if I -- I'm not sure if, in my mind's eye, I have the same agreement you're referring to.

Q. Do you know whether Nalin Mishra was an onsite consultant to RELX?

A. My recollection is that Nalin Mishra was an Informatica employee who was present on -- at RELX to help with the Informatica software.

Q. And is it your belief as part of those duties, he actually installed hardware -- or added hardware to the platform?

A. My recollection is that he stated -- that is Mr. Mishra stated in e-mails that he brought up that he's added certain servers to the ICCE platform that included Informatica software.

Q. And can you identify those e-mails?

A. As I sit here right now, I don't remember exactly which they were, but I'm happy to identify those at some later point.

Q. But as you sit here now, you don't know?

A. I don't remember exactly which e-mails they are. I do remember reading e-mails to that effect where there was an e-mail thread about when certain servers would be added, and Mr. Mishra was involved in that thread, and at some point said, you know, we encountered some difficulties, but now they're -- they're online and part of the ICCE platform in the production environment.

Q. And could you tell from those e-mails whether Mr. Mishra was merely putting software on the server or whether he was connecting up the servers to the platform?

A. Probably want to review the e-mails to be sure. He may have had a role in the software installation or the -- or adding the -- or put it this way, or reconfiguring the hardware so that it was part of the ICCE platform or some combination of the two. I'm not 100 percent sure as I sit here right now.

Q. And your support for that statement is these e-mails?

A. That's my recollection of the e-mails, yes.

Q. And do you know whether -- what were the roles based on those e-mails of Wisvari and Li? Do you remember?

A. Based on those e-mails, I don't recall, as I sit here right.

Q. Do you believe that Mr. Mishra also purchased the servers?

A. I don't remember, as I sit here right now.

Q. Okay. So on December 5, 2014, RELX had 72 licenses; is that correct?

A. That is my recollection.

Q. Yet you point to this sizing model which says only what? How many? Six CPUs are necessary?

A. As produced, that is the maximum number that it suggests.

Q. Is that mentioned in this set of e-mails?

A. Well, the file is there.

Q. I asked if it's in the e-mails.

A. Well, I think the attachments are part of the e-mails, so yes.

Q. Okay. Disregard the attachment for a second. Is it in the e-mails? Let me help you out.

A. Well, I see a -- this is -- the Bates number INFA 0000218569. This is an e-mail from Christopher

Boytim. It says about halfway through this e-mail: "To determine if we go forward with an" -- and then in parentheses for each of these numbers -- "6, 8, 12 CPU core system." So I don't know exactly the context for this, but at least the number 6 does appear here with respect to trying to determine perhaps how many CPU cores would be appropriate.

Q. And does it say that Mike Tomechak and Gil owe you both a capacity recommendation?

A. That is -- those words are present in this e-mail, yes.

Q. Okay. Did you talk to Gil Rosen or Mike Tomechak about this?

A. No, I didn't speak to them at all.

Q. And is this sizing consistent with the information that you used in the sizing model?

A. Well, in this e-mail, for instance, the number 6 is referenced with respect to how many C -- or how many cores each system would have. That's also the number that is the maximum in the document. So there does appear to be some relation at least.

Q. Do you know why the sizing model and capacity were being discussed?

A. I don't recall, as I sit here right now.

Q. Do you know if this is for a new project?
A. I'm not sure about the overall context to these e-mails. They do mention PowerCenter, for instance, however.
Q. But it's your understanding, because you use the sizing model in your opinions, that it was for the existing ICCE platform with Informatica on it at that time?
A. That was an assumption I made in my report, yes.
Q. If that assumption is wrong, how does that impact your report?
A. Well, if the document I looked at as part of the report, if there was evidence to show that it was related to a different project and not -- and completely unrelated from the ICCE platform, than -- or rather then -- then I don't think this document would -- would speak to planning for the ICCE platform, assuming there were compelling evidence that it was instead designed to be used for a different project using completely different Informatica software.
Q. So that would mean your opinions are incorrect in your report?
A. Well, I'm saying if we -- if we assume

different things, then my opinions would change.

Q. Sir, is it correct in your conversation with Mr. Vellturo that you talked about the access of alternative -- I'm sorry. Strike that.

In your conversation with Mr. Vellturo, did you tell him about the ability to use alternatives to the Informatica software?

A. We did discuss alternatives that RELX could have used instead of the Informatica software.

Q. And what'd you say about that?

A. Well, my recollection is in particular I mentioned using Java and Perl, as RELX had done in the past to perform a similar operation that the Informatica software was designed to perform.

Q. And do you know whether or not the Informatica software had any advantages over Java and Perl?

A. We didn't discuss that in any detail during the conversation that we had. I imagine there would be pros and cons to both approaches.

Q. Are you aware that there's documentation that talks about advantages of the Informatica software over Java and Perl?

A. I don't remember seeing specific documentation to that effect, but it could exist, I suppose.

Q. Would that change your opinions at all?

A. Well, I'd have to consider the totality of the benefits and disadvantages of using certain software in certain circumstances. But even if there were alternatives that were better or worse for certain reasons, it seems like they would still be alternatives. There would just be a calculus that would have to be performed by whoever was making the decision about which avenue to pursue.

Q. Are you aware of whether -- yeah. Sure, there could be -- still be alternatives, but they could -- the alternatives could -- you know, could present issues such as higher maintenance costs, lower processing speed, things like that, right?

MR. SCOTT: Objection. Calls for speculation.

A. I suppose in theory they might, but I think you would have to consider all the pros and cons of both -- or all the alternatives that were available. I don't think it's -- or I'd have to consider all the information available with respect to the different dimensions along which certain alternatives may be better or worse than others.

Q. Sure. Did you do that?

A. Well, in my conversation with Chris Vellturo,

we discussed what other alternatives might have been viable. We didn't go into detail about the pros and cons of all the alternatives.

Q. Do you know what he said about alternatives in his report?

A. I have not reviewed his report.

MR. DOYLE: All right. Let's mark the deposition of Jeffrey on this case as Exhibit -- 8?

(Exhibit No. 7, Jeffrey Reihl deposition transcript marked for identification)

Q. Could you go to page 35, please, sir. First of all, do you know who Jeff Reihl is?

A. One moment. I assume it's in his deposition here somewhere, but I understand he's employed by RELX in some capacity.

Q. Well, do you know who he is? I'm just asking.

A. Well, I haven't spoken to him, but I forget his title, as I sit here right now.

Q. Do you understand that he's the chef technology officer at RELX?

MR. SCOTT: Objection.

A. As I sit here right now, I'm not sure, but I believe the CTO of RELX was deposed at some point, so I'm sure it's in this deposition somewhere, and if you

represent that to me, I'll accept it for now.

Q. Could you go to page 35.

A. Okay. I'm there.

Q. Do you see where it says: "Were there any alternatives that were brought to your attention instead of using the Informatica software?" That was the question in the deposition.

Mr. Riehl answered: "There were no other alternatives that I recall that were brought forward." Do you see that?

A. Well, let me just read the couple lines ahead and before.

Q. Sure.

A. So in his response, he says: "I am very confident that the team looked at different options including writing the software ourselves, and likely felt that the Informatica software would help us accelerate time to market and was a tool that we could utilize to do that."

So it sounds like there were other options that were considered, including writing the software themselves.

Q. Sure. But which one did they select?

A. My understanding is that they decided to

purchase licenses for the Informatica software.

Q. But he does say: "There are no other alternatives that I recall that were brought forward." Do you see that?

A. I do. And then he further clarifies that he's confident that the team looked at different options as well.

Q. Sure. But he's answering that none were brought forward to him, the CTO, right?

A. That may be what he's saying. He doesn't specify to whom they were brought forward.

Q. Did you ask him about it?

A. Sorry. Did who ask whom about what?

Q. Did you ask Jeff Reihl about what he said here?

MR. SCOTT: Objection. Asked and answered.

A. I didn't have any conversations with Jeff Reihl.

Q. I mean, in your infinite wisdom, you said there were a lot of alternatives to the Informatica software, in your opinion, right?

A. Well, I did mention that there were other avenues through which ICCE -- I'm sorry -- RELX could have accomplished what -- what they had set out

to do with the ICCE platform.

Q. Did you say anything else other than that?

A. I'm sorry. Did I say anything else?

Q. Yes.

A. I don't remember the complete deposition testimony that I've given today.

Q. No, I'm not asking about the deposition testimony. When you spoke to Vellturo, what exactly did you tell him about alternatives?

A. Well, we discussed alternatives such as the Java and Perl scripts that RELX had used in the past, and those may have been a viable way forward with the ICCE platform.

Q. Do you know if those alternatives worked as well as the Informatica software?

A. I didn't do any detailed analysis about the pros and cons of those approaches. It's my understanding that that was something that was considered.

Q. Did you discuss alternatives with anybody at RELX?

A. As I sit here right now, I'm not sure, but if I did, it would have been with Mr. Groff probably.

Q. But you can't recall right now as you sit here?

A. Well, I'm not certain, but I think it may have come up when we were talking about the specific Java and Perl scripts that were still utilized as part of the ICCE platform even with the Informatica software incorporated.

Q. And do you know whether any of the alternatives had all of the performance advantages that the Informatica software had?

A. Well, I'm not sure which performance advantages Informatica software had --

Q. Why not?

A. -- in general.

Q. Why not?

A. Well, we looked at some documents that it wasn't clear what those benefits were, whether they were aspirational. It's hard for me to answer the question without a knowledge of what the general benefits of Informatica software were, if there were any. And that wasn't something I focused on in my reports.

Q. If that's the case, how can you say that they're alternatives?

A. Well, just because I don't have a detailed understanding of all of the pros and cons of all the alternatives, I think I can still say that alternatives

existed and that such alternatives -- at least my understanding is they were considered to a degree.

Q. I don't think we can say you have a detailed understanding of all the pros and cons, or that you even have a non-detailed understanding of pros and cons.

MR. SCOTT: Objection. Mischaracterization --

Q. What --

MR. SCOTT: -- of testimony.

Q. -- advantages existed with respect to the Informatica platform?

MR. SCOTT: Objection. Asked and answered.

Q. Can you name any?

A. When you say "advantages," advantages compared to what?

Q. Compared to anything.

A. Well, I think there were some aspirational benefits that were recorded in documents we reviewed earlier today. I don't know or haven't done analysis to show that those benefits actually manifested themselves in any way.

Q. Why not?

A. One of the primary opinions in my report was examining the question of the degree to which RELX may have benefited from having the ICCE platform have access

to, for instance, 104 cores versus a smaller number of cores for which they were licensed. And so the question of that differential benefit is different from whether there was a benefit overall to the Informatica software, and so I didn't examine it in much detail.

Q. What factors impact computer performance, in your mind?

A. In general, some examples, in a non-exhaustive list, would include CPU speed, the amount of memory that's installed on the computer, the speed of I/O access to the hard drive or flash drive that's present on the computer, the networking speed of the computer to the Internet or to -- or to other computers.

There might be other factors related to specific software configurations such as you could connect it to a VPN, a virtual private network, that may slow down network communications in some degree in exchange for further benefits. There are probably other considerations, but there are a lot of factors with respect to how a computer might perform in various circumstances.

Q. And does the type of workload matter when answering the question?

A. So specifically to the Informatica software,

does the type of workload matter? With respect to performance? Is that your question?

Q. I had asked about computer performance in general.

A. Well, in general's different from with respect to the Informatica software, but let me think about that for a moment.

Q. Well, strike that.

What steps are involved in capacity planning?

MR. SCOTT: Objection. Asked and answered.

A. Well, in general, with capacity planning, the objective is to determine what sort of computer system you might need to perform certain tasks, and so the things you might consider are some of the things that I mentioned earlier like CPU speed, network speed or bandwidth, hard drive access speed, among other items.

Q. Do you need to measure current use?

A. When you say "current use," what are you referring to?

Q. Current use of the processing power or processors.

A. Do you need to consider it? Well, it depends on what you're planning for. If you're planning for something that is the same or similar to what you're

doing at the current time, then you might find that information useful. I don't think you necessarily need to consider it, but it depends on the circumstances for the exercises -- the exercise that you're performing at the time.

Q. What if you're planning for growth?

A. Again, it would depend if the growth were similar to what you're doing at the current time or the same or if it's different. Growth in general can be growth along a number of different dimensions, so I'd want to know more about the specific hypothetical that you have in mind.

Q. And how would you measure current use?

A. Well, if we're talking about CPU utilization, then you may look at -- may look at data that you have for CPU utilization for the computers that you have deployed at the current moment.

Q. Do you know of anything else other than CPU utilization in terms of what you would measure to measure current use?

A. Assuming again that you're talking about current use of CPUs, in terms of capacity planning, you may want to consider how many servers that you -- that you currently have or may want to have. You may want to

look at specifically what those cores and those servers are -- are doing. It could be that your -- like your current setup has servers that have cores that are doing processes that you're -- or executing processes that are unrelated to what you're planning for, and so you may want to consider that as well.

Q. How would you collect the data for those types of measurements?

A. Well, one thing you should do if it were a Linux system is you could run a UNIX command called "top" which shows you the different processes that are -- that are executing on the computer, and so you could see, for instance -- you know, come to some conclusion about what percentage of CPU utilization was related to system processes that the operating system needs versus the processes that you care about for planning or other processes that may be unrelated to the tasks that you're looking to perform.

Q. And when software is running, it uses resources behind -- besides the CPU, right?

A. When software's running, it can use resources like memory or disc or network bandwidth that would be different from the CPU. There are -- there are others as well.

Q. Would you take -- collect data for those measurements as well of those different devices?

A. For the different resources. If you expected that some of these resources might be a bottleneck, it may be interesting to look at the degree to which those resources are currently being utilized, if you expect the tasks that you're planning for to be similar or the same to the tasks that the system is performing now.

Q. Would you collect information on peak utilization?

A. It all depends on what you're planning for. If peak utilization matters to you for some reason over a certain time period, it may be interesting to consider the degree to which that was -- those peak utilization events were occurring on the existing system.

Q. What metrics would you use those measurements of peak utilization?

A. Would you --

Q. What metrics would you use in your determination of peak utilization?

A. It depends on what you're trying to measure. Peak utilization would probably be defined along dimensions of CPU utilization, and then also the amount of time that that CPU utilization occurred over. So it

would depend on your needs for the specific environment and what you -- what you cared about in terms of peak utilization.

Q. You had mentioned bottlenecks before. How would you identify a bottleneck?

A. Well, generally, you would look for whether there were certain resources that were close to maximum capacity for whatever resource that -- or the resources that you're considering. So if you're -- if you're monitoring the system currently and all of the memory's being used and you might be thrashing because parts of the program are being swapped out fairly consistently, then memory might be something that you look at.

But in general, you're looking for whether there are resources that are -- that are close to maximum capacity, and so you might want to increase capacity for those resource first before you examine others.

Q. Did you determine -- did you attempt to determine what the bottlenecks are in the ICCE system using Informatica?

MR. SCOTT: Objection. Assumes facts not in evidence.

A. So specifically to the question of the degree to which RELX may have benefited from having cores

available to the -- or, for instance, 104 cores available for the ICCE platform and incorporated Informatica software, because the -- the issue was CPU capacity measured in cores in this case, that was the item that I focused on.

It's possible there were other bottlenecks to the system such that if you were to increase the number of CPU cores for a certain configuration, then those cores would provide no benefit to you because the system would be bottlenecked on some other resource.

Q. How can a bottleneck in one area, for example, I/O, affect utilization of other resources like the CPU?

A. Well, in general, if you have a computer system and it is bottlenecked on I/O, then that is the sort of determining factor in terms of how much processing can be performed by the computer system. So if your -- your system is maxing out on -- in accessing the hard drive in that way, if you were to increase other resources of the system like the number of CPUs or the number of CPU cores or memory, if it's truly a bottleneck, then you wouldn't see any difference in performance because the system would still be bottlenecked on that same I/O resource.

Q. Do you know what the RETS team was at RELX?

A. As I sit here right now, I don't remember what the acronym stood for.

Q. So you don't know what their role was at RELX?

A. As I sit here right now, I don't remember.

Q. Did you talk to anyone from RETS?

A. Well, I talked to Mr. Groff and Mr. Hoffman, so if they're part of the RETS team, then I did, but I'm not sure if they were.

Q. Did you ask whether they do capacity planning in RETS?

A. I don't think I asked that specific question.

Q. Did you ask anyone that was involved in doing capacity planning for the ICCE network using Informatica at any time during the period?

A. Would you repeat that? I'm confused about "ICCE network."

Q. Okay. All I'm asking is did you ever talk to anyone at RELX who actually did capacity planning on the ICCE network or for the ICCE network?

A. Well, I talked to Mr. Groff and Mr. Hoffman. So if they performed capacity planning, then I would have talked to -- I talked to them. But I don't know -- I don't know whether they did.

Q. So you don't know whether or not you talked to

anyone that actually did capacity planning in building the network, the ICCE network; is that right?

MR. SCOTT: Objection.

A. Well, with respect to the ICCE platform, I only talked to Mr. Groff and Mr. Hoffman, so if they performed capacity planning, then I did speak to individuals at RELX that did that; and if they didn't, then I didn't because those were the only individuals that I spoke to from RELX.

Q. Did -- did the folks at RELX, anybody, model for future capacity needs as it related to ICCE and also the Informatica software on the ICCE platform?

MR. SCOTT: Objection. Calls for speculation.

A. I didn't talk to anyone who -- who mentioned doing that, so as I sit here right now, I'm not aware of anyone who did that.

Q. Do you know if RELX did capacity planning for ICCE?

A. As I sit here right now, I don't recall a specific example of when they -- they would have done that.

Q. I didn't ask you when they would have done it. I'm just asking do you know if RELX did capacity planning for ICCE?

A. So as I sit here right now, I don't recall an instance where -- where I talked to someone who mentioned that RELX was doing capacity planning.

THE VIDEOGRAPHER: We have reached seven hours.

MR. SCOTT: Let's take a short break, and then we'll let you know if we have any redirect.

MR. DOYLE: Okay. Will you give me five more minutes?

MR. SCOTT: No.

THE VIDEOGRAPHER: The time is 7:34 p.m. we are off the record.

(Whereupon the deposition was concluded at 7:34 p.m.)

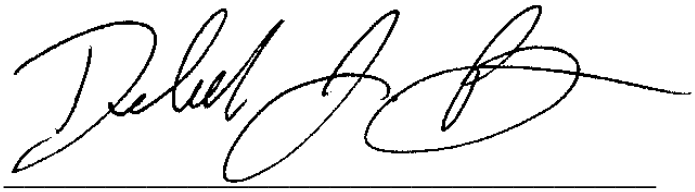
COMMONWEALTH OF MASSACHUSETTS

ESSEX COUNTY

I, DEBORAH J. BATEMAN, Court Reporter and Notary Public in and for the Commonwealth of Massachusetts, do hereby certify that the witness whose deposition is hereinbefore set forth, was duly sworn and that such deposition is a true record of the testimony given by the witness.

I further certify that I am neither related to or employed by any of the parties in or counsel to this action, nor am I financially interested in the outcome of this action.

I witness whereof, I have set my hand and seal this 3rd day of July, 2018.



Deborah J. Bateman, Notary Public in and
for The Commonwealth of Massachusetts
My Commission Expires: November 2, 2023

DEPOSITION ERRATA SHEET

Our Assignment No. J2333077

Case Caption: RELX INC. vs INFORMATICA LLC

DECLARATION UNDER PENALTY OF PERJURY

I declare under penalty of perjury that I have read the entire transcript of my Deposition taken in the captioned matter or the same has been read to me, and the same is true and accurate, save and except for changes and/or corrections, if any, as indicated by me on the DEPOSITION ERRATA SHEET hereof, with the understanding that I offer these changes as if still under oath.

Signed on the _____ day of _____, 20___.

_____

CHRISTOPHER RUCINSKI

DEPOSITION ERRATA SHEET

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

SIGNATURE:_____DATE:_____

CHRISTOPHER RUCINSKI

DEPOSITION ERRATA SHEET

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

Page No._____Line No._____Change to:_____

_____

Reason for change:_____

SIGNATURE:_____DATE:_____

CHRISTOPHER RUCINSKI